

Processing Framework and Match-up Database

MODIS Algorithm

Version 2

by

Robert H. Evans
University of Miami
Miami, FL 33149-1098

OCTOBER, 1996

TABLE OF CONTENTS

PREFACE.....	IV
1.0 INTRODUCTION	1
1.1 Algorithm and Product Identification.....	1
1.2 Algorithm Overview.....	1
1.3 Document Scope.....	2
1.4 Applicable Documents and Publications	2
2.0 OVERVIEW AND BACKGROUND INFORMATION	2
2.1 Experimental Objective.....	2
2.2 Historical Perspective.....	3
3.0 DESCRIPTION OF ALGORITHM.....	3
3.1 Introduction based on AVHRR-Oceans Pathfinder Matchup Databases	3
3.1.1 Global matchup databases	4
3.1.1.1 SST Matchup Database	5
3.1.1.2 Time Coordinates.....	7
3.1.1.3 Generation of a satellite data extraction list	8
3.1.1.4 Extraction and subsetting Level 1a fields	9
3.1.1.5 Matchup of the <i>in situ</i> and satellite quantities	12
3.1.1.6 Quality control and cloud flagging.....	12
3.1.1.7 Ocean color matchup database.....	13
3.1.1.8 Quantities in the oceancolor matchup record.....	14
3.1.1.9 Merging for the Matchup process.....	17
3.1.2.1 Cloud filtering in the AVHRR Matchup database	17
3.1.2.2 MPFSST Coefficient estimation.....	24
3.1.2.3 Transition to MODIS	25
3.1.3 Transition of ocean color sensors: processing and calibration	26
3.1.3.1 Introduction	26
3.1.3.2 Relevant Properties of Case 1 Waters.....	29
3.1.3.4 Retrospective Characterization --- Methods.....	33
3.1.4 Variance or uncertainty estimates.....	1
3.1.4.1 Observation frequency for useful <i>In situ</i> - satellite matchups.....	1
3.1.4.3 Initialization and Long-Term Degradation.....	5
3.1.5 Retrospective Characterization -- CZCS Results	8
3.1.5.1 Long-term Variation	8
3.1.5.2 Band 4	9

3.1.5.3 Bands 2 and 3	10
3.1.5.4 Band 1	10
3.1.5.4 Short-term Fluctuations	10
3.1.6 Concluding Remarks	12
3.2 Practical Considerations	13
3.2.1 Overview	13
3.2.2 Programming/Procedural Considerations	14
3.2.2.1 Overview	14
3.2.2.2 SST and ocean color processing requirements	14
3.2.2.4 Overall ocean processing time for the Level-2 radiances	19
3.2.2.5 Data Volume	20
3.2.2.5.1 Volume of data to be transferred to Miami	23
3.2.2.5.2 Total volume	23
3.2.2.6 Network Capacity	23
3.2.3 Calibration Validation	23
3.2.3.1 Introduction	23
3.2.3.2 Satellite Field Analysis Techniques	24
3.2.3.3 Temporal comparison of temperature fields	28
3.2.4 Quality Control and Diagnostics	30
3.2.4.1 Introduction	30
3.2.4.2 Potential quality effects of stray light	32
3.2.4.3 Quality control of <i>in situ</i> observations	34
3.2.4.4 Ocean Processing Level-2 Quality control flags	37
3.2.4.4.1 SST atmospheric correction processing flags	38
3.2.4.4.2 Additional SST Flags To Be Added	38
3.2.4.4.3 Ocean color atmospheric correction processing flags	40
3.2.5 Exception Handling	45
3.2.6 Data Dependencies	46
3.2.7 Output Product	46
4.0 CONSTRAINTS, LIMITATIONS, ASSUMPTIONS	46
5.0 REFERENCES	47
APPENDIX 1 EQUAL-AREA GRIDDING SCHEME	1
A1.1 Introduction	1
A1.2 Overview	1
A1.3 The poles	4
A1.4 Binning software	4
APPENDIX 2 DATA DAY DEFINITION	1
A 2.1 INTRODUCTION	1
A2.2 A 24-hour data day	2
A2.3 A spatial data-day definition	4
A2.4 How is the beginning of a data-day defined?	6

A2.5 Advantages of the Spatial Definition of a Data-Day	9
A2.6 Other Issues.....	14

LIST OF TABLES

Table 1. Continuous Pathfinder time coordinate.....	8
Table 2. Quantities included in the distributed Pathfinder matchup databases.....	10
Table 3. In situ quantities tentatively identified for inclusion in the Ocean color matchup database.....	15
Table 4. Satellite quantities tentatively identified as needed for ocean color algorithm development.....	16
Table 5. Quantities to be extracted at the time of Level1	17
Table 6. Validation test.....	24
Table 7. Bias and Slope for CZCS.....	27
Table 8. Quantities and auxiliary data required for atmospheric correction.....	30
Table 9. Buoy retrieval statistics for fixed mooring and NOAA-9 AVHRR.....	1
Table 10. CZCS Initialization Constant	8
Table 11. Processing loads for SST and ocean color.....	15
Table 12. File sizes for SST and ocean color products.....	20
Table A2-1. Beginning times of fifteen data-days for descending orbits, NOAA-11 spacecraft.....	9

TABLE OF FIGURES

Figure 1. Flow diagram of the “Filter 1” tests.....	19
Figure 2. Flow diagram of the “Filter 2” tests.....	20
Figure 3. Prune classification tree NOAA-9.....	23
Figure 4a. Normalized water-leaving radiance at 440 nm; as a function of pigment concentration.....	29
Figure 4b. Normalized water-leaving radiance at 520 nm as a function of pigment concentration.....	30
Figure 4c. Normalized water-leaving radiance at 550 nm as a function of pigment concentration.....	30
Figure 5a. Global frequency distribution of normalized water-leaving radiance at 520 nm for a 10-day period in mid June 1981.....	35
Figure 5b. Global frequency distribution of normalized water-leaving radiance at 550 nm for a 10-day period in mid June 1981.....	35
Figure 5c. Global frequency distribution of normalized water-leaving radiance at 520 nm for a 10-day period in early October 1981.....	35
Figure 5d. Global frequency distribution of normalized water-leaving radiance at 550 nm for a 10-day period in early October 1981.....	35
Figure 6. Long-term degradation of the sensitivity $f=g-1$ of the 4 CZCS.....	9
Figure 7. Calibration variation in the bands at 520nm (solid line) and 550 nm (dotted line).....	11
Figure 8. Pathfinder SST image for 1991, week 12.....	26
Figure 9. Reynolds blended analysis for 1991, week 12.....	27
Figure 10. Pathfinder - Reynolds difference map.....	28
Figure 11. Four year presentation of Pathfinder-Reynolds difference for zonally averaged.....	29
Figure 12. Same presentation as Figure 11 for week.....	30
Figure 13. Pathfinder channel 3 - channel 4.....	33
Figure 14. OA Pathfinder SST - Buoy SST for all buoys for 1992 to 1994	34
Figure 15. Mean OA Pathfinder SST - Buoy SST per buoy	35
Figure 16. Plot of Buoy and OA Pathfinder SST	36
Figure 17. Plot of Buoy and OA Pathfinder SST.	37
Figure A1-1. Number of 9.28 km tiles per zonal row as a function of latitude (North or South).....	2
Figure A2-1. Descending NOAA-11 tracks for a 24-hour data-day	2
Figure A2-2. Locations of the boundaries of 24-hour data days for a 10-day period.....	4
Figure A2-3. NOAA-11 descending orbits for a spatially-defined data-day beginning on July 26 1992,.....	5
Figure A2-4. AVHRR nadir track and scan lines for a 20-minute.....	6
Figure A2-5. Latitude of crossing of the 180° meridian for NOAA-11 descending orbits.....	8
Figure A2-6. NOAA-11 descending orbits for spatially-defined data-day.	11
Figure A2-7. Data-day beginning on July 27 1992 15:14:00 UTC.....	12

Preface

This Algorithm Theoretical Basis Document (ATBD) describes our current working model for construction of a matchup data base for calibration of the ocean visible and infrared bands of the MODIS sensor and integration of the ocean product algorithms into a cooperative group of programs. While effort has been made to make this document as complete as possible, the reader should understand that this version of the document is a snapshot of ongoing work, *i.e.*, the algorithm development is an evolving process.

As will be seen from reading the document, there are areas that still require research effort before finalization. In particular this effort depends on continued availability of algorithm descriptions and codes from the participating ocean investigators.

The results described in this document are based on ongoing joint development and tests associated with the NASA/NOAA Pathfinder AVHRR Oceans activity and SeaWiFS projects. Experience gained with these efforts is directly assisting development of the MODIS comparison database with respect to design, testing and implementation.

1.0 Introduction

Document focuses on implementation of a Level-2 framework incorporating the Level-2 algorithms of the MODIS Ocean Team investigators (Abbott, Brown, Carder, Clark, Esaias, Evans, Gordon, Hoge, Parslow). The implementation first addresses the conversion of L1 counts into calibrated radiance or reflectance, conversion of these at satellite radiances into brightness temperatures or water leaving radiances L_W , and generation of the products by executing the individual product algorithms.

1.1 Algorithm and Product Identification

The calibration data set produced by the algorithm will be labeled version 2. This is a level 2 product with EOSdis product number 3303; it is MODIS product number 34, labeled Calibration Data.

1.2 Algorithm Overview

This algorithm is being developed on the MODIS Ocean Team Computing Facility (MOTCF) for use in the EOS Data and Information System (EOSdis) core processing system. The individual ocean product algorithms described in the respective MODIS Ocean Team PI ATBD's will be combined into a structure that provides easy integration into the MODIS TLCF leading to final porting into the EOSdis environment.

Initial focus is Level-2 conversion of Level-1 counts/reflectance into water leaving reflectance for the visible and brightness temperatures for the infrared. Algorithms described in the ATBD's for Gordon and Brown are used to describe considerations affecting the development of appropriate match-up data bases, integration of their algorithms into a Level-2 program and subsequent comparison of the algorithm results with the in-situ observations. Considerations for generating Level-3 products from the Level-2 are presented in Appendices 1 and 2. The Level-3 fields provide an important validation component of the algorithms integrated in the Level-2 program.

1.3 Document Scope

This document describes the basis for the Calibration database, gives the structure of the current version 1 algorithm, discusses implementation dependencies on other observing streams, and describes validation needs.

1.4 Applicable Documents and Publications

MODIS Proposal, 1990, Processing and Calibration for Ocean Observation with EOS/MODIS, Robert H. Evans

MODIS Execution Phase Proposal, 1991, Processing and Calibration for Ocean Observation with EOS/MODIS, Robert H. Evans

2.0 Overview and Background Information

2.1 Experimental Objective

The goals of this proposal are two fold: generation of a matchup database that will serve as a basis for algorithm validation and integration of algorithms developed by the MODIS Ocean Team (MOT) into a coherent processing package. Validation of the water leaving radiances will be effected through application of a match-up database that includes appropriate satellite and *in-situ* measurements. Initial uses of the matchup data base include testing of color and SST algorithm performance as a function of viewing geometry (satellite and solar zenith angle), space (latitude and longitude), time, a variety of geophysical conditions (e.g. windspeed, ozone, water vapor concentration) and state of the sensor (e.g. sensor temperature, mirror position, orbit). Previous applications of this approach include characterizing CZCS long term sensor drift and development and evaluation of candidate SST algorithms for the ocean Pathfinder project.

Generation of the matchup database and integration of the MODIS ocean team algorithms utilizes the SCF developed at the University of Miami/RSMAS. The facility presently includes DEC ALPHA workstations and a multi-processor SGI computer. Sufficient processing power is available to generate MODIS ocean products at full resolution for a reasonable fraction of the world's oceans. However, rather than processing the full resolution data set, a subset of the data will be analyzed to test calibration and validate algorithms.

2.2 Historical Perspective

Heritage programs provide the basis for MODIS algorithm. A program developed for the Pathfinder ocean SST product forms the framework for analyzing AVHRR derived SST, algorithm development and validation, and application of the matchup database. Development for ocean color algorithms is based on experience gained in transition from CZCS to SeaWiFS algorithms. Development of the SeaWiFS program involves integration of algorithms generated by H. Gordon (atmospheric correction) and K. Carder (chlorophyll) with additional MODIS algorithms provided by F. Hoge, D. Clark and M. Abbott. In addition, data validity tests are being developed through collaboration of the SeaWiFS CAL/VAL team and our group. The SeaWiFS and Pathfinder programs have been converted to the MODIS implementations and coded using the FORTRAN 90 language and EOS coding standards.

3.0 Description of Algorithm

3.1 Introduction based on AVHRR-Oceans Pathfinder Matchup Databases

Algorithm development supported for this effort includes generation of a matchup database (MDB) that consists of spatially and temporally matched satellite and *in situ* observations and application of this database to assist in MOT algorithm development and validation. The following sections describe current efforts where a MDB has been created for NOAA AVHRR instruments and used to analyze the behavior of AVHRR SST retrieval equations.

As part of the Oceans component of the AVHRR Pathfinder project sponsored by NOAA and NASA, a "matchup" data bases are compiled, combining *in situ* surface temperature (SST) measurements (and, in some cases, other *in situ* environmental variables) and AVHRR data corresponding to the times/locations of *in situ* measurements. The purpose of the matchup data bases is the development, testing and validation of Pathfinder SST algorithms.

Extension to include population of an *in situ* portion of the ocean color database will become possible in late Fall, 1996 following the initial test deployment of the MOBY instrument by D. Clark.

3.1.1 Global matchup databases

Experience in producing and evaluating the matchup database (MDB) has been gained using the NOAA AVHRR 4-km Global Area Coverage (GAC) observations. The database has been applied in the following areas: generation of SST retrieval equation coefficients for the AVHRR instruments, evaluation of spatial and temporal trends, quality control of the *in situ* validation data set and comparison with radiative transfer simulations.

The matchup databases for MODIS are based on procedures developed for the AVHRR MDB. Spatial resolution of the database will be 1-km rather than the 4-km available with AVHRR GAC observations. The initial MODIS MDB might be forced to continue using the present 4-km resolution as a fall back option should either a limitation in subsetting capability or data delivery capability impact access to the global Level-1 MODIS observations.

3.1.1.1 MODIS Matchup Databases

Two matchup database products; MOD18_L2_Bouy and MOD28_L2_Bouy, will be developed for the MODIS sensor. These databases will contain temporally co-located *in situ* measurements and subsetted Level-1a and Level-2 satellite data for ocean color and sea surface temperature, respectively. Each of these databases contains both a "real-time" and a retrospective component.

The "real-time" component is a preliminary database containing *in situ* records that were available electronically on the day that the level-1 granule is processed at the DAAC. A month of daily "real-time" matchups are assembled into a single file and should be available for distribution shortly after the close of each month. This component of the database is relatively small and typically contains moored buoy locations representing only 20% of *in situ* records collected globally on a given day. The limited spatial coverage of the "real-time" component requires a second, retrospective approach.

The retrospective component contains *in situ* data sources from both moored and fixed buoys, and for ocean color, shipboard measurements which are gathered over a 1 year time period from sources not available in "real-time". The corresponding year's archived Level-1a data is then searched to extract the satellite information. A year of retrospective matchups are split into individual months and should be available for distribution within the first quarter of the following year. Details of the processing and factors which may affect the timing of the distribution of these products are presented below.

Compilation of the Matchup databases:

There are four main common steps involved in the compilation of the Matchup databases:

- *Obtaining and reformatting of *in situ* data
- *Extraction of the subsetting satellite quantities
- *Matchup of the *in situ* and satellite quantities
- *Quality control and flagging of cloud contaminated data

These common aspects will first be discussed in the context of the SST Matchup database. Differences associated with the Ocean color Matchup database will be discussed separately in a later section.

3.1.1.2 SST Matchup Database

Obtaining and Reformatting of *in situ* data:

The SST *in situ* data originate from two main sources, moored and drifting buoys. The *in situ* sources for the MODIS sensor will be the same as those presently used in the Pathfinder Oceans Matchup Database(PFMDB). The agencies supplying the data are:

Moored buoys:

US National Data Center (NDBC)

Japan Meteorological Agency

TOGA/TAO Project office(NOAA Pacific Marine and Environmental Laboratory)

Drifting buoys:

AOML(NOAA Atlantic and Oceanographic and Meteorological Laboratory)

MEDS(Canadian Marine and Environmental Data Service)

The retrospective component of the database obtains data from the above sources via FTP or tape directly from the archival agency after the data has been subjected to various stages of quality control and made available to the general science community. In 1995 200,000 *in situ* records compiled from these sources were used in the AVHRR MDB. These agencies have been continually expanding their network of buoys over the last few years. If this trend continues we estimate that 300,000 records per year will be contained in the retrospective component of the Sea Surface temperature Matchup Database for the MODIS sensor.

Many of the agencies post a portion of the daily data collected from their buoy network on the GTS system for real-time distribution. The US. Navy (NAVOCEANO) monitors the GTS and daily assembles a file containing data posted by all of the above sources. The Navy file is then released and electronically transferred to Miami RSMAS the following morning. In 1995 the retrospective and real-time components of the AVHRR MDB contained 772 and 571 unique buoy IDs, respectively. It is important to note that in the retrospective component a buoy reports 6-8 times a day. In contrast, buoys in the real-time component report 1-2 a day.

In situ data compiled from all of the sources are reformatted into a standard format and very general quality control is applied to the data. The quality control includes testing that the position is not on land, SST value is within a reasonable range, and valid date and time values are present. Drifting buoys present in the retrospective component are separated by unique ID's and the speed and acceleration of each buoy is checked. The

buoys in the real-time component are continually monitored by unique ID for the same quality issues. If a buoy is found to have failed the quality control tests from the preceding day it will be removed from the satellite extraction list for the incoming day.

After passing quality control tests, the original platform IDs (usually WMO or ARGOS ID) are modified to facilitate the identification of the various *in situ* SST sources once all the matchups are assembled by adding a source dependent constant to the original ID. Buoys IDs in the real-time component are not modified and remain as the WMO ID. To facilitate the matchup process of the *in situ* and satellite records the dates and times of the *in situ* SST reports are converted to a continuous time coordinate. The Gregorian dates and UTC time of the *in situ* data are included in the Matchup Database for convenience of the users.

3.1.1.3 Time Coordinates

To facilitate the matchup process, dates and times of both the satellite and *in situ* data are converted to a continuous time coordinate, “seconds since January 1, 1981,” here referred to as “Pathfinder seconds”. Table 1 shows the values of Pathfinder seconds for the beginning of each month in 1982-89. The values correspond to day 1 of each month at 00:00:00 GMT. These values can be subtracted and then the actual date can be obtained through a few simple calculations.

Table 1. Continuous Pathfinder time coordinate values for the beginning of each month, 1986-1989.

	1986	1987	1988	1989
Jan	157,766,400	189,302,400	220,838,400	252,460,800
Feb	160,444,800	191,980,800	223,516,800	255,139,200
Mar	162,864,000	194,400,000	226,022,400	257,558,400
Apr	165,542,400	197,078,400	228,700,800	260,236,800
May	168,134,400	199,670,400	231,292,800	262,828,800
Jun	170,812,800	202,348,800	233,971,200	265,507,200
Jul	173,404,800	204,940,800	236,563,200	268,099,200
Aug	176,083,200	207,619,200	239,241,600	270,777,600
Sep	178,761,600	210,297,600	241,920,000	273,456,000
Oct	181,353,600	212,889,600	244,512,000	276,048,000
Nov	184,032,000	215,568,000	247,190,400	278,726,400
Dec	186,624,000	218,160,000	249,782,400	281,318,400

3.1.1.4 Generation of a satellite data extraction list

Due to the time consuming nature of the satellite data extraction process for the retrospective component of the database, we have developed a procedure to exclude those times and locations for which there was not a satellite pass within ± 30 minutes from the *in situ* measurements. The procedure involves the computation of the Time of Closest Approach (TCAP) with respect to a given point and time. This procedure relies on orbital routines provided by D. Baldwin (Univ. of Colorado). The TCAP filter significantly reduces the time required for the extraction process. The TCAP filter outputs records containing buoy ID, latitude, longitude, time, orbit #, and slant distance for *in situ* records which fall within the position/time window. These records are then assembled into individual day extraction lists. *In situ* records which occur within -30 minutes of the start of day N or +30 minutes of the end of day N are placed on the extraction list of day N. For MODIS, the time criterion will match the MODIS time to the closest available buoy time, typically this will be from the 'morning' NOAA satellite.

The building of extraction lists for the real-time component is different than that of the retrospective component due to several constraints. The satellite data subsetting and extractions will be done in near real-time with the Level 1 processing thereby eliminating the lengthy archive search time present in the retrospective component. In addition, the total number of *in situ* records available at the time of Level-1 processing is significantly less. The decrease in both the volume of data and search time precludes the need for TCAP filtering. A timing constraint still remains in regard to the receipt of the *in situ* data used to build extraction lists and the processing of Level1 granules. The

real-time *in situ* file for day N is received on day N+1. Typically an *in situ* file for day N contains records for a 3-4 day time window centered on day N. In the AVHRR MDB this constraint is handled by filling the extraction list for day N with records from *in situ* files received for day N to N+7 which fall on day N. The extraction list for day N is submitted for extraction on day N+7. For the Pathfinder AVHRR, we are able to hold a rolling 7-10 day temporary archive of GAC data on-line to facilitate extraction. We assume that the on line storage of MODIS Level 1 granules will limit completeness of the real-time extraction list. We are planning to fill the extraction list with position/time from *in situ* drifting and fixed buoy records received at the time of required extraction list close. The extraction list for day N will then be supplemented with known fixed buoy positions which we expect to receive for that data day after list close. Analysis of the Navy *in situ* files shows that 80 and 90% of the *in situ* records posted to the GTS are received within 4 and 7 days, respectively of *in situ* collection.

The compilation of all of the *in situ* data and the construction of extraction lists will occur at the University of Miami RSMAS. The actual extraction of the satellite information will occur at the DAAC or potentially at another facility such as the MODIS Team Leader Computer facility. A less desirable alternative will be to extract the satellite records using subsampled (4x4) Level 1 granules at the Miami SCF. A single extraction list file will be submitted daily for the real-time component. The retrospective component will submit 365 extraction list files once annually after all *in situ* data sources for the given year have been received.

3.1.1.5 Extraction and subsetting Level 1a fields

Satellite data from the level 1a granules will be extracted for various size boxes centered at each *in situ* location. Once the granule to be extracted has been identified, the geographical position of the *in situ* point must be inverted to granule coordinates of scan line and pixel. The initial extraction subset includes the level-1a counts, information needed to convert the level 1a to level-1b, pixel geometry, and other ancillary data. For SST this includes the level 1a pixel counts, cloud mask, reference climatology, geolocation, aerosol, and calibration files. It is anticipated that 2,000 positions will be extracted a day.

The extraction box size for 25 positions will be 1000*1000 pixels, 75 positions at 512*512 pixels, and 1900 at 100*100 pixel boxes. The larger sized boxes will correspond to special long term study sites where it will be desirable to follow evolution of ocean or atmospheric features in the area..

The initial extraction subset is then processed to level-1b (brightness temperatures) and other quantities needed for algorithm development and evaluation. Summary statistics will then be determined for a 5x5 pixel box centered on the *in situ* position. The summary statistics for each of the relevant channels (bands 20, 22, 23, 26, 31 and 32) of

the satellite record includes the central pixel value, mean and median of all values inside the box, and the minimum and maximum values within the box. Also included in the satellite record will be information relevant to sensor calibration, and the slopes and intercepts for the counts-to-radiance conversions. Using the geometric information, a sun glint index is calculated to assist in the identification of pixels contaminated by glint. An example of the fields included in the satellite record of the present AVHRR MDB is shown in Table 2.

Table 2. Quantities included in the distributed Pathfinder matchup databases. The variable names given in the last column (“Code”) are used in the “Record filtering and flagging” section.

Field	Field Description	Units	Code
1	<i>In situ</i> date (Gregorian, YYMMDD)	yr,mon,day	yymodd
2	<i>In situ</i> time (UTC, HHMMSS)	hr,min,sec	hhmmss
3	Satellite time (Pathfinder coordinates)	seconds	stime
4	Latitude of center pixel	Decimal deg.	slat
5	Longitude of center pixel	Decimal deg.	slon
6	Average PRT temperature	°C	prt
7	Solar zenith angle	Degrees	solz
8	Satellite zenith angle	Degrees	satz
9	Sun glint index	sr ⁻¹	glnt
10	Emissivity , channel 3	—	em3
11	Emissivity , channel 4	—	em4
12	Emissivity , channel 5	—	em5
13	Central value of 5x5 pixel box, channel 1	—	ch1
14	Central value of 5x5 pixel box, channel 2	—	ch2
15	Central value of 5x5 pixel box, channel 3	°C	ch3
16	Central value of 5x5 pixel box, channel 4	°C	ch4
17	Central value of 5x5 pixel box, channel 5	°C	ch5
18	Median of 5x5 pixel box, channel 1	—	med1
19	Median of 5x5 pixel box, channel 2	—	med2
20	Median of 5x5 pixel box, channel 3	°C	med3
21	Median of 5x5 pixel box, channel 4	°C	med4
22	Median of 5x5 pixel box, channel 5	°C	med5
23	Minimum value of 5x5 pixel box, channel 1	—	min1
24	Minimum value of 5x5 pixel box, channel 2	—	min2
25	Minimum value of 5x5 pixel box, channel 3	°C	min3
26	Minimum value of 5x5 pixel box, channel 4	°C	min4

Field	Field Description	Units	Code
27	Minimum value of 5x5 pixel box, channel 5	°C	min5
28	Maximum value of 5x5 pixel box, channel 1	—	max1
29	Maximum value of 5x5 pixel box, channel 2	—	max2
30	Maximum value of 5x5 pixel box, channel 3	°C	max3
31	Maximum value of 5x5 pixel box, channel 4	°C	max4
32	Maximum value of 5x5 pixel box, channel 5	°C	max5
33	Average value of 5x5 pixel box, channel 1	—	av1
34	Average value of 5x5 pixel box, channel 2	—	av2
35	Average value of 5x5 pixel box, channel 3	°C	av3
36	Average value of 5x5 pixel box, channel 4	°C	av4
37	Average value of 5x5 pixel box, channel 5	°C	av5
38	PRT 1 Temperature	°C	prt1
39	PRT 2 Temperature	°C	prt2
40	PRT 3 Temperature	°C	prt3
41	PRT 4 Temperature	°C	prt4
42	Central value channel 1 (same as field 13)	—	cm1
43	Central value channel 2 (same as field 14)	—	cm2
44	Central value w/ emissivity correction, ch. 3	°C	cm3
45	Central value w/ emissivity correction, ch. 4	°C	cm4
46	Central value w/ emissivity correction, ch. 5	°C	cm5
47	Calibration slope, channel 3	$\text{mW m}^{-2} \text{ cm}^{-1}$ $\text{sr}^{-1} \text{ count}^{-1}$	slope3
48	Calibration slope, channel 4	see Field 47	slope4
49	Calibration slope, channel 5	see Field 47	slope5
50	Calibration intercept, channel 3	$\text{mW m}^{-2} \text{ cm}^{-1}$ sr^{-1}	intcp3
51	Calibration intercept, channel 4	see Field 50	intcp4
52	Calibration intercept, channel 5	see Field 50	intcp5
53	Time of <i>in situ</i> SST, Pathfinder coordinates	Seconds	btime
54	Buoy latitude	Decimal deg.	blat
55	Buoy longitude	Decimal deg.	blon
56	Buoy ID	—	bid
57	<i>In situ</i> SST	°C	bsst
58	Delta-SST (First-guess sat. SST minus <i>in situ</i> SST)	°C	sst1
59	Filter code (1 or 2)	—	pass

In the AVHRR MDB only the summary statistics for the original subset level-1a processed to level-1b 5x5 pixel box is retained in the satellite portion of the record. For the sensor it may be desirable to make the original subset level-1b pixel data available and include a file pointer in the matchup database. This would enable end users of the database to derive contextual information.

3.1.1.6 Matchup of the *in situ* and satellite quantities

The next step in the generation of the matchup databases is to temporally match *in situ* records against satellite extractions. To limit the variability introduced by the time separation between the two data sources (Minnett 1991), the absolute difference between the time of the *in situ* report and the time at which that location was viewed by the satellite (matchup time window) is restricted to a maximum of 30 minutes or to the closest available *in situ* record where these data are obtained via another satellite. *In situ* records that do not fall within the stipulated time window will be rejected. Furthermore, the real-time satellite records may have data for predicted *in situ* times which were not received, these will also be rejected. *In situ* records which pass the temporal matchup must subsequently pass a spatial test. A maximum distance of 0.1° in latitude and longitude will be allowed between the *in situ* location and the location of the central pixel in the extraction box. The matchup procedure for TOGA/TAO buoys will be slightly different from those used for other data sources. These buoys differ from other sources because the reported SST values represent averages over longer time periods. The reported SSTs for the TOGA/TAO buoys are the average of six measurements taken every 10 minutes; the reporting time is the end of each averaging period. The *in situ* time for the center of the matchup window and reported in the matchup database will correspond to the center of the *in situ* averaging period.

The output of the matchup process is a series of records which contain both satellite-derived and *in situ* derived data. In a small number of cases, the satellite part of the matchup record is the same for two *in situ* records collected a short time apart. For instance, some NDBC buoys report data every 30 minutes. In these cases the same satellite extraction may satisfy the matchup window for two consecutive *in situ* observations. Only the matchup record with the closest *in situ* and satellite time will be retained.

3.1.1.7 Quality control and cloud flagging

A large number of the matchup records will have obvious problems such as gross cloud contamination and should not be used for algorithm development. A series of cloud tests will be developed by the University of Miami RSMAS to identify high quality matchups records which can be confidently used in algorithm development and

coefficient estimation. Presently we are planning to distribute all matchup records and include a quality flag indicating if the record passed the Miami cloud tests. This will enable end users the ability to develop other cloud tests if desired. A description of the present development of new cloud filters and coefficient estimation by the University of Miami for the Pathfinder program can be found in Appendix A. The final stage of the database is to assemble matchup records into individual monthly files which will be delivered to the DAAC for distribution. The cloud flags developed by Menzel will be included once their procedures are validated.

3.1.1.8 Ocean color matchup database

The overall scheme of the process used to create the ocean color matchup database; MOD18_buoy will be the same as that for the SST matchup database described above. The main difference between the two databases is the source and structure of the ocean color *in situ* data and the number of satellite quantities to be included in the final matchup record. The *in situ* data compiled from various sources include drifter and moored buoys, and shipboard along-track and profile measurements of bio-optical quantities. In contrast to the SST measurements, the *in situ* bio-optical quantities in the database will have both a vertical and horizontal measurement for a given geographical location when available. This three dimensional nature of the ocean color data will require that a matchup record for a given time and geographical location be separated into multiple files which share at least one field in common so as to link the *in situ* information. The matchup files submitted to the DAAC for distribution will be flat files. The University of Miami will create a relational database for these files which will allow custom queries to create unique views of the quantities present across the matchup records. The team members and the general bio-optical scientific community is presently being solicited for input on the *in situ* and satellite quantities which should be included in this database. Sources and quantities which have been tentatively identified are discussed below.

Data sources:

The real-time component will consist of data from the MOBY moored buoy and the MOCE cruises located off Hawaii being delivered by Dennis Clark and associates. Several other bio-optical moored buoys are planned by a number of different Agencies and these will be added as they become available. We are estimating that 20 moored and drifting buoys will eventually be available real-time.

The retrospective component will consist of *in situ* data available from the calibration-validation effort and other Agencies which include bio-optical measurements in their respective field programs. Sources tentatively identified include the JGOFS field program, Bermuda Atlantic Time Series (BATS), Hawaii Ocean Times series (HOTS), and the Japanese YBOM in support of OCTS. Cruise data from individual investigator program will be included as they are identified. In the event that the SEABASS-

SIMBIOS bio-optical database in support of SEAWIFS is still available it will also be included as an *in situ* data source.

3.1.2.9 Quantities in the oceancolor matchup record

In situ matchup record:

Table 3 shows the *in situ* quantities which have been identified from the MOCE datasets for inclusion in matchup record. Many of these quantities are measured by several different instruments using different wavebands and depth resolutions. Where appropriate the matchup record will contain the same quantity measured or derived by several different techniques or instruments. For example the water leaving radiance (Lw) in the MOCE dataset is measured by both the Biospherical MER and the MOS-SIS high resolution spectrometer. Data from both instruments will be included in the matchup record. A field with instrument and technique flags will need to be added for many of the quantities.

Table 3. *In situ* quantities tentatively identified for inclusion in the Ocean color matchup database. Many of these quantities consist of both a depth and spectral measurement.

<i>In situ</i> quantity	code
date	date
time	time
latitude	lat
longitude	lon
depth	Z
water- leaving radiance	Lw
normalized water- leaving radiance	nLw
diffuse attenuation downwelling	Kd
diffuse attenuation upwelling	Ku
optical depth	1/Kd
reflectance	R
remote sensing reflectance	R1
Photosynthetically available radiation	PAR
transmission	%t
beam attenuation coefficient	c
chlorophyll fluorescence	ChlFl
aerosol optical thickness	AOT
chlorophyll concentration (fluoremetric)	chl
chlorophyll concentration (hplc)	chla-hplc
Chlorophyll fluorsecence	chl _f
Phycobilin pigment concentration	PUB,PEB
Coccolith concentration	cocco
total particulate absorption	aph
detrital absorption	dph
dissolved organic matter concentration	DOC
total suspended solids	ses
air temperature	airT
windspeed	ws
relative humidity	%rh
seasurface temperature	sst

instrument/technique flags	flag
----------------------------	------

Satellite matchup record:

The satellite portion of the matchup record will include the statistical summary for the level 1b data derived from the level-1a extraction subset for all of the visible channels and any other channels, calibration, and ancillary data identified for algorithm development. In addition the level 1b subset will be processed to level-2 and a statistical summary of the level 2 products will be included. It is not planned to include all 36 level-2 products in the matchup record. Only those level-2 products (i.e.. Lw's + others) identified as required for algorithm development will be present. Other intermediate quantities used or derived during Level-2 processing which are important in algorithm development may also be include. A list of satellite quantities presently identified as important for algorithm development is presented in Table 4.

Table 4. Satellite quantities tentatively identified as needed for ocean color algorithm development.

Satellite quantities	Processing level
pixel counts	level 1a (bands 8-16 ,5-7,26)
calibration information	level 1b (bands 8-16 , 5-7,26)
pixel geometry	level 1b (bands 8-16 , 5-7,26)
reflectance	level 1b (bands 8-16, 5-7,26)
aerosol reflectance (La)	-
aerosol optical thickness (tau)	-
aerosol model used	-
Lw	level 2 (bands 8-16, 5-7,26)
chlorophyll concentration	level 2
Coccolith concentration	level 2
Fluorescence line height	level 2
Dissolved organic matter	level 2
Phycobin concentrations	level 2
seasurface temperature	level 2
various quality flags	Level 1 and 2

Ancillary matchup record fields:

The addition of the level-2 data in the ocean color matchup record requires that the level-1a extraction and subsetting process include all information needed as input to the level-1b and level-2 processing PGE's. The data types required as output of the matchup extraction and subsetting process is shown in Table 5. As suggested in the SST matchup

database, it may be desirable to make the level 1b and level-2 subsetting pixel data available for contextual purposes and include a file pointer in the matchup record.

Table 5. Quantities to be extracted at the time of Level1 subsetting to allow subset processing to level1b and level 2 products.

Subsetting products and quantities
Level 1a raw pixel counts MOD01
calibration information Level1b
geolocation MOD03-L1a
cloud mask MOD06
aerosol MOD 04
3 week L2 ocean color reference field
SST MOD28
Ancillary meteorology and ozone

3.1.1.10 Merging for the Matchup process

The merging the ocean color *in situ* and satellite records in the matchup process will be the same as described above for SST. The matchup record will include the satellite data *in situ* measurements, and a quality flag indicating if the record passed our cloud flagging technique. It is planned that a single monthly matchup file will be delivered to the DAAC in an HDF format. However due to the complexity of the *in situ* record multiple files may be required.

3.1.2 MATHEMATICAL ASPECTS OF THE ALGORITHM

3.1.2.1 Cloud filtering in the AVHRR Matchup database

In the AVHRR MDB a large number of records have serious cloud contamination. For that reason initially a sequence of 15 filters were implemented to exclude the most cloud-contaminated matchups. The first set of filters (Figure 1) excludes from the database any of the records which fail any of the tests thus eliminating gross cloud contamination. Records which pass the first set of filters are then passed through a second set of filters (Figure 2). A record which fails any of the "Filter 2" tests is then flagged for possible cloud contamination. Many of the "Filter 2" tests are similar to those in "Filter1", however thresholds are more restrictive. Generally only 10% of the total matchup records were able to pass these two filter sets. While records passing were cloud free, a large portion of "good" matchups were also being rejected. We have

therefore begun investigating other techniques to identify cloud contaminated matchups.

Figure 1. Flow diagram of the “Filter 1” tests. Matchups that fail any of the tests in this suite are excluded from the Pathfinder matchup data base. The variable names used are listed in Table 3.

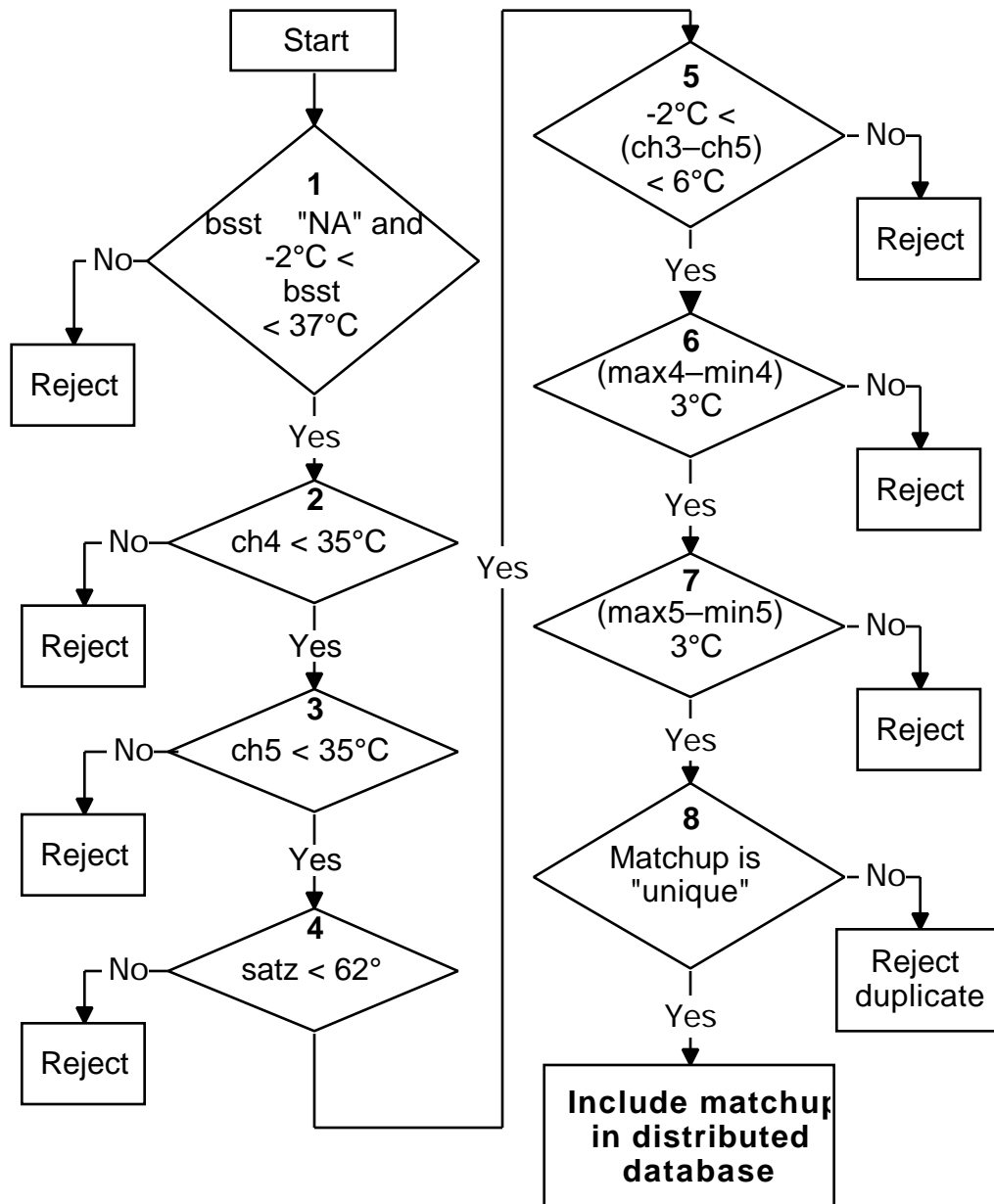
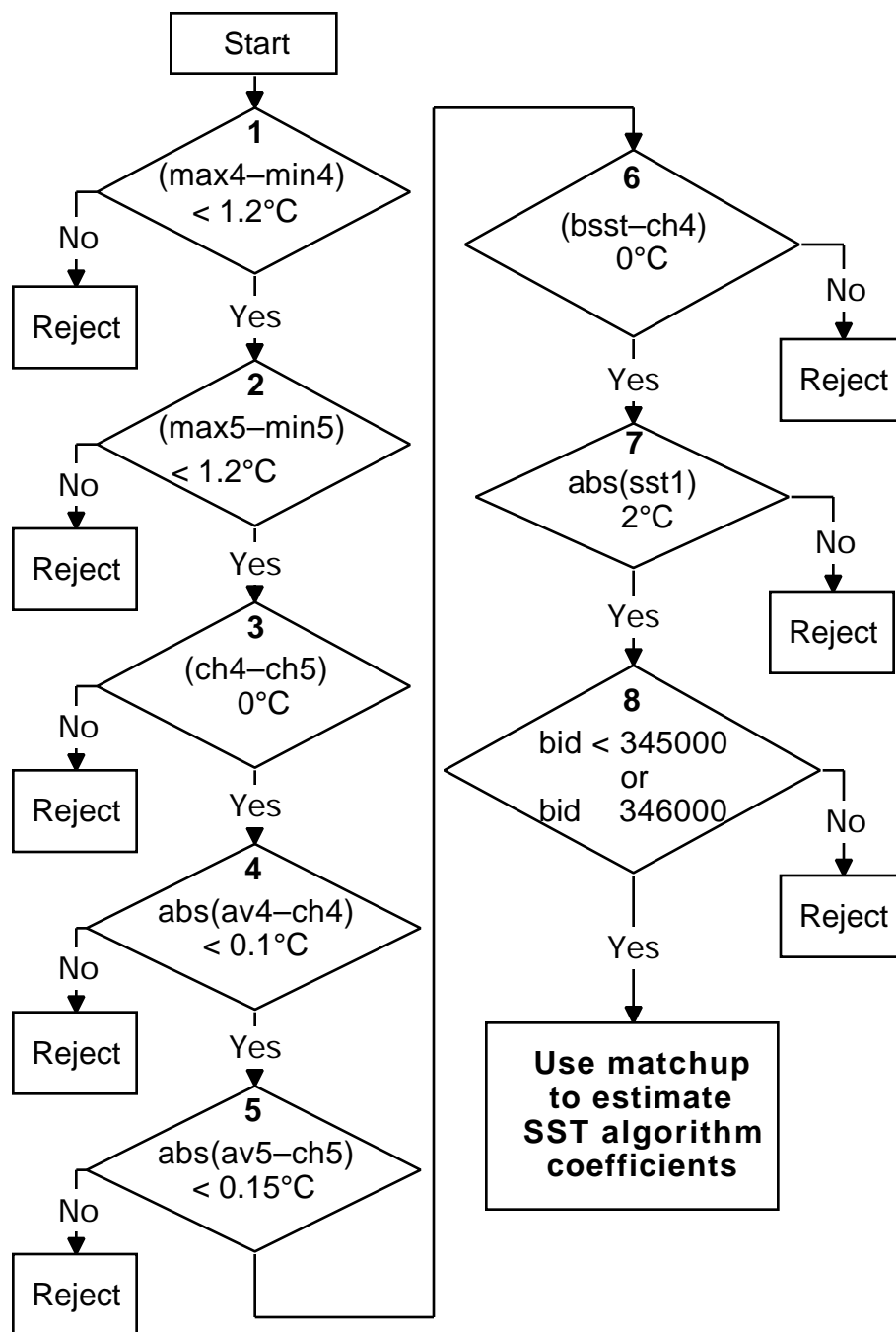


Figure 2. Flow diagram of the “Filter 2” tests. Matchups that fail any of the tests in this suite are not excluded, but the result of the tests is encoded in field 59 of the matchup data base. The variable names used are listed in Table 3. See text of this section for a definition of “sst1”.



A promising technique is the use of tree based models to classify matchup records. Tree-based models provide an alternative to linear logistic and additive logistic models for classification problems (L.A. Clark and D. Pregibon,1991). A classification tree is a collection of many rules determined by binary recursive partitioning. Tree based models are developed by successively splitting the training dataset into increasingly homogeneous subsets.

The process of growing a Tree based model for use in classifying cloud-free ("good") and cloud-contaminated ("bad") matchup records involves several steps.

1. Initial filtering of PFMDB for quality control (reasonableness).
2. Define category membership of record as "good" or "bad".
3. Creation of training and validation datasets.
4. Selection of variables to be used in the Tree model.

Initial filters:

These tests contain thresholds for the channel brightness temperatures to ensure that the satellite data had reasonable geophysical values and thereby remove records which contain digitization errors. Our experience has been that biased final satellite SST values are obtained from the central pixel located in non-homogeneous extraction boxes. We therefore included a homogeneity test (max-min <0.7 of 5x5 extraction box)) for both channels 4 and 5.

Define category membership:

Category membership of a record was determined from the residual of the buoy SST minus a First Guess SST. The First Guess SST may be calculated using operational algorithms (i.e. NOAA-NESDIS NLSST) or coefficients developed from other cloud flagging routines such as described earlier. A "good" record is defined as having an abs(residual) $\leq 2^0$, records while an abs(residual) $> 2^0$ is defined as "bad". The category membership is dependent on having an unbiased estimate of the First Guess SST. The median of the "good" residuals must be close to zero to use the First Guess SST for membership definition. If the algorithm used for the First Guess SST results in a biased estimate, alternative formulations for the First Guess SST must be found.

Training and Validation sets:

The training set was created by randomly selecting 1/3 of the night time records from the matchup database. The selection process included a probability weighting function so that the training set would contain approximately equal numbers of both "good" and "bad" records. The use of only nighttime records was chosen to minimize the possible bias in threshold values for the splits as a result of solar contamination. The goal was to create a generalized tree which could be used to classify both day and night records in regards to cloud contamination only. The training set was used to develop the Tree based model locating the important binary splits. The remaining 2/3 of the nighttime

records were used as the validation set. The validation set provided independent data on which to verify the generality and misclassification rate of the tree model developed from the training set.

Tree variables:

Spectral quantities which have been previously shown to indicate cloud contamination were used as variables in the tree based model. The spectral variables used in the tree model were homogeneity (max-min) for AVHRR channels 3, 4, and 5, channel difference for ch3-ch5, ch4-ch3, and ch4-ch5, and linear model values for channels 3 and 5 which correspond to MODIS channels 20, 31 and 32. The limitations of the software (Splus) to identify splits based on optimal linear combinations of variables required the pre-calculation of these linear models using least square regression techniques. The pre-calculations were done for channel 3 as a function of channel 4, channel 3 as a function of channel 5, and channel 5 as a function of channel 4. The coefficients for each of the functions were determined by least square regression using the channel data from all nighttime matchup records defined as having a "good" membership. Finally, the satellite zenith angle was collapsed into 4 groups (0-29, 30-39, 40-49, >50 degrees) and was also included as a variable in Tree estimation.

Tree estimation:

The training set was then recursively partitioned on the above 10 variables and allowed to run to completion. The complete tree for NOAA-9 had 89 terminal nodes and was probably over fitted to the training set. Recall that the recursive binary partitioning technique attempts to make the membership of each terminal node as homogeneous as possible. To limit this over fitting of the tree model to the training dataset, the full tree model was then pruned (Figure 3) to create a more parsimonious tree but still retain a reasonable misclassification rate (8%). The pruned tree model was then used to classify the records in the Validation set. The tree predicted classification for the Validation set was then compared to the actual membership (Table 6). The results from the validation test verify that the pruned tree model was not over fitted to the training set and the misclassification rate remained low. The next step was to repeat verification process using daytime records. This test demonstrated that the final tree model was robust with a misclassification rate of 8.8%. This tree model can then be used to classify NOAA-9 matchup records of unknown quality. Matchup records classified by the tree as "good" are then used in coefficient estimation for the Miami Pathfinder SST algorithm (MPFSST).

Figure 3. Prune classification tree NOAA-9. Boxes represent the terminal nodes of the tree. Number of records classified in each terminal node is shown. The purity of the records contained in the node is shown in parenthesis. Symbols used in the splits represent brightness temperatures; T35= ch3-5, t45=ch3-ch5, T3=ch3.

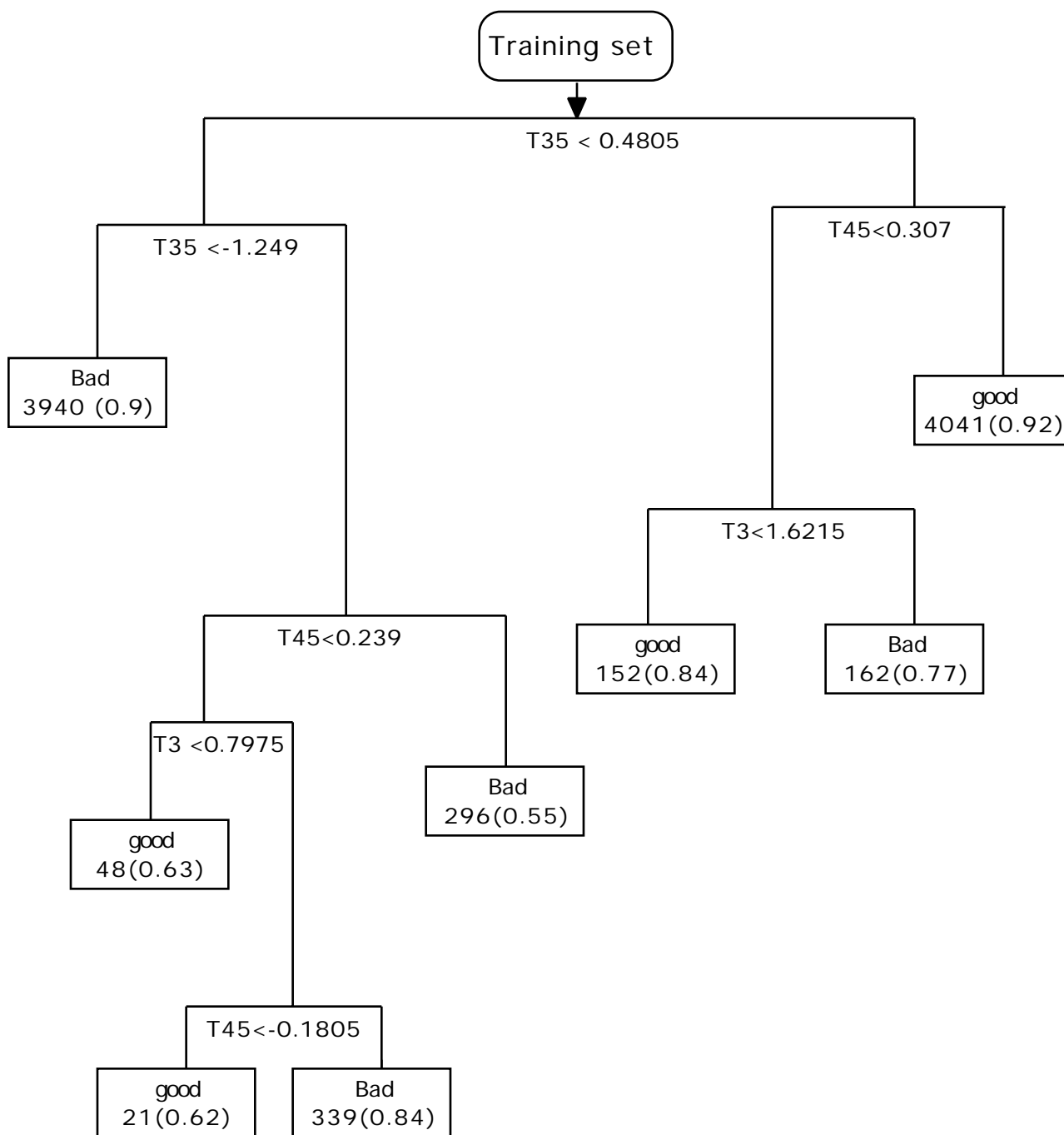


Table 6. Validation test. Cross tabulation of actual classification of record from residuals and predicted classification from pruned tree model. Total misclassification rate is 13% for the validation dataset.

classification category	predicted Bad	predicted good	actual classification
Bad	2522	252	2775
Good	1089	6411	7500
Total	3611	6411	10275

3.1.2.2 MPFSST Coefficient estimation

Coefficients are determined for a two piece algorithm based on the different atmospheric regimes. The regime is defined by the channel 4-5 difference (T45). A set of coefficients are determined for each month (N) using matchup records from a 5 month (N-2 to N+2) moving window. Coefficients are estimated by an iterative process involving robust linear regression techniques. Coefficients are determined in a 3 steps process.

1. Use robust regression procedures to determine initial coefficient values for the SST equation.
2. Define robustness weights
3. Perform weighted least squared regression procedures based on the robustness weight.

Robust regression procedure:

A robust regression procedure was used to determine the initial coefficients for the Pathfinder SST equation.

$$\text{Pathfinder SST} = C1 + (C2 * \text{ch4}) + C3 * (\text{ch4} - \text{ch5}) * \text{bsst} + C4 * (\text{ch4} - \text{ch5}) * (1 / \cos(\text{satz})) - 1$$

Where C1:C4 = regression coefficients, bsst = buoy SST, ch= channel brightness temperature, satz= satellite zenith angle.

A robust regression is made using matchup records from each atmospheric regime which were classified as "good" by the cloud flagging routine. The robust procedure is less sensitive to remaining outliers such as those caused by cloud contamination and errors in the buoy SST. The residuals from the robust regression are then used to determine weights for the subsequent least square regression.

Robustness Weights:

The weights used for the least squared regression were the product of the residual and monthly time weighting factors. Residual weights are determined by applying a Bi-square function to the residuals of the robust regression. The cut off for the Bi-squared function was set at 6*MAD (median of the absolute value of the residuals). The monthly time weighting factors are set at 0.5 , 0.8, 1.0, 0.8, 0.5 for months N-2 to N+2 when estimating coefficients for month N.

Weighted Least square linear regression:

Coefficients are then estimated for each T45 regime (ch4-ch5 <0.7 and ch4-ch5 > 0.7) for month N. The weighted least squared regression uses the matchup records from the 5 month time window centered on month N and classified as being in the given T45 regime. Only records with weights greater than zero are included in the regression.

3.1.2.3 Transition to MODIS

The MODIS calibration activity presented in this document builds on the work by the MCST. We expect to receive information on the relative calibration of the detectors for each band, calibration dependency on scan angle and the absolute calibration. Additional information that influences calibration is detector linearity and spectral stability of the filters together with platform and sensor parameters such as navigation, pointing and temperature. The algorithms and tables necessary to normalize outputs of within band detectors and correct to mirror position produced by the MCST will be used in this effort. The prime point of departure will be application of the final calibration coefficients that this algorithm utilizes to produce absolute calibration. We will utilize the vicarious method to adjust the coefficients to achieve agreement between the atmospherically corrected $[L_w(i)]_N$ described in the next section and the corresponding *in situ*. observations. In particular, we will utilize the results obtained from MCST study of the MODIS solar calibrations to test trends obtained from the vicarious method.

Method of application:

Get Level-1 counts. Apply MCST algorithms for normalizing detectors, linearity, mirror correction (scan angle, mirror side, polarization), time correction (sensor drift, temperature). It is expected that most of these corrections will utilize lookup tables. The final conversion of counts to radiance will utilize correction coefficients determined from use of the vicarious method.

The initial version of these tables will be based on pre-launch thermal vacuum test data. After launch methodology tested and validated in the SeaWiFS program will be utilized to update MODIS pre-launch coefficients. This sensor initialization is described in Gordon's ATBD. Finally, if the SeaWiFS, ADEOS/OCTS or COLOR sensors are operational, we will be able to compare retrievals from each of these sensors.

These steps are briefly summarized below and defined more fully in the following section: 1)Apply pre-launch calibration data. 2)Obtain radiances, use satellite location and pointing information. 3)Locate satellite observation corresponding to *in-situ* observation. 4)Apply quality control, cloud test procedures. 5a)Visible - Use Gordon's sensor initialization procedure to obtain initial set of calibration coefficients. 5b)Infrared - Use thermal vacuum test data to obtain counts to radiance conversion. 6)Apply appropriate atmospheric correction procedure to obtain geophysical quantity (e.g. radiance, temperature). 7)Compare satellite and *in-situ* quantities as a function of observation parameters. Adjust calibration coefficients to obtain agreement.

3.1.3 Transition of ocean color sensors: processing and calibration

3.1.3.1 Introduction

The planned approach for calibration of the SeaWiFS and MODIS sensors is based on experience gained from calibrating and processing eight years of CZCS data; the heritage 8 bit, four channel ocean color instrument. The scheme outline below is presented in Evans and Gordon [1993]. A simple functional form for converting the observed counts into radiance $L_t(i)$ in Band i is given by

$$L_t(i) = S(i, G) \times DC + I(i, G) \quad (3)$$

where DC is the digital output of the instrument (0-255), and $S(i, G)$ and $I(i, G)$, respectively, are called the slope and intercept. The radiometric sensitivity of the instrument in Bands 1-4 was adjustable on command from the ground. This was achieved by varying the Gain (G) of the amplifiers between the detectors (one for each band) and the digitizer. Four Gains were available ($G=1, 2, 3$, and 4), and the (nominal) prelaunch values of $S(i, G)$ and $I(i, G)$ for each Gain determined by radiometric calibration in the laboratory are reproduced in Table 7 [Ball Aerospace Division, Boulder CO, 1979]. Equivalent tables for the upcoming ocean color instruments will be generated using data obtained by the instrument calibration team during thermal vacuum test. Due to the increased precision available with SeaWiFS (10 bit) and

MODIS (12 bit), band or possible detector level calibration will likely be represented as a series of tables specifying the counts to radiance transform.

Calibration requirements for space-borne sensors viewing the ocean are significantly more demanding than for land-viewing sensors. This is due to the fact that the albedo of the ocean is very low and the atmosphere contributes a major portion (80-100%) of the radiance measured at the sensor. Therefore, if the water-leaving radiance is to be determined with an absolute accuracy of 10%, the error in the sensor calibration cannot exceed ~ 1%. Since the desired calibration error is somewhat less than that which can be realized even under ideal laboratory conditions, it is necessary to calibrate the sensor in orbit. This is accomplished by adjusting the calibration to force the sensor *plus* algorithms to yield ship-measured water-leaving radiances *in situations* where the latter were available [Gordon ,1987]. This is called "vicarious" calibration and is similar to methods used for other space-borne radiometers [Koepeke, 1982; Fraser and Kaufman, 1986, Slater et al., 1987]. Of course, vicarious calibration also is required to assess any variation in the sensitivity of the instrument with time.

Table 7. Bias and Slope for CZCS; $S(i, G)$ and $I(i, G)$ $mW / cm^2 \mu mSr$ for the four CZCS Gains, $I(i, G)^*$ are revised values of $I(i, G)$ as described in the text.

Band	$S(i, G)$	$I(i, G)$	$I(i, G)^*$	Gain
1	0.04452	0.03963	0.03963	1
2	0.03103	0.06361	0.05361	1
3	0.02467	0.07992	0.08013	1
4	0.01136	0.01136	0.01136	1
1	0.03598	0.05276	0.03963	2
2	0.02493	0.08826	0.06361	2
3	0.02015	0.06247	0.08013	2
4	0.00897	0.03587	0.01136	2
1	0.02968	0.02879	0.03963	3
2	0.02032	0.09752	0.06461	3
3	0.01643	0.06570	0.09503	3
4	0.00741	0.02963	0.01136	3
1	0.02113	0.03359	0.03963	4
2	0.01486	0.05647	0.06361	4
3	0.01181	0.04723	0.09159	4
4	0.00535	0.01604	0.01136	4

The vicarious calibration effort for CZCS began shortly after launch when it was realized that inconsistencies appeared during application of the atmospheric correction algorithm [Gordon, 1981a; Gordon, 1981b]. Comparison of predicted radiance at the top of the atmosphere (based on atmospheric and in-water measurements) with that measured by CZCS led to suggested adjustments in the pre-launch sensor calibration [Viollier, 1982; Gordon et al., 1983]. Thus, it was necessary to modify Eq.(3) to include an additional factor:

$$L_t(i) = k(i, G) \times S(i, G) \times DC + I(i, G) \quad (4)$$

where the constant $k(i, G)$ is now referred to as the "initialization" constant. In $k(i, G)$, G refers to the gain setting of the instrument, and is included because of a small change in k with the sensor gain. It is the factor relating the prelaunch calibration to the in-orbit calibration. At about the same time that the need for $k(i, G)$ was established, it was observed that the sensitivity of the instrument was decreasing with time. This required a further modification of the calibration equation:

$$L_t(i) = g(i, t) \times k(i, G) \times S(i, G) \times DC + I(i, G) \quad (5)$$

where the factor $g(i, t)$ accounts for the loss of sensitivity with time, t . The main difficulty with effecting vicarious calibration is that subsurface and atmospheric optical measurements are required. These subsurface optical measurements were available only during the initial field experiments conducted for validation, i.e., to determine how well the system --sensor plus algorithms-- could provide the phytoplankton pigment concentration, C , which is defined to be the sum of the concentrations of chlorophyll a and its degradation product phaeophytin a . The SeaWiFS mission provides a realistic testbed for MODIS since the ocean color bands share similar characteristics, e.g. spectral range, and the *in situ* observation capabilities (ship and buoy) that will be utilized for MODIS are being deployed for SeaWiFS (see ATBD by Clark). Thus SeaWiFS provides a multi-year experiment to optimize procedures and integrate algorithms prior to the first MODIS flight opportunity.

In contrast with the *in situ* program planned for SeaWiFS and MODIS, very few optical cruise data sets and no data from optical moorings were available during the CZCS lifetime. This led to long temporal gaps between times when calibration could be achieved by adjusting $g(i, t)$ to force the closest agreement possible between the ship-measured and CZCS-retrieved pigment concentrations. If this procedure is followed frequently under a variety of atmospheric conditions (low and high aerosol content), the resulting set of calibration constants will be close to the true values. For CZCS, the procedure was applied whenever high quality surface measurements were available

and this enabled initial documentation of the long-term sensitivity variation. However, even accounting for the sensor degradation, artifacts, *i.e.*, large unexplained errors in the retrievals, still occurred. This led us to believe that short-term fluctuations in the sensitivity must have been occurring, *i.e.*, higher frequency sensitivity variations superimposed on the (assumed) smooth degradation in sensitivity.

3.1.3.2 Relevant Properties of Case 1 Waters

At this point it is useful to summarize the relevant optical properties of Case 1 waters, *i.e.*, waters for which phytoplankton and their immediate detrital material, along with the water itself, control the optical properties of the medium. Case 1 waters play a

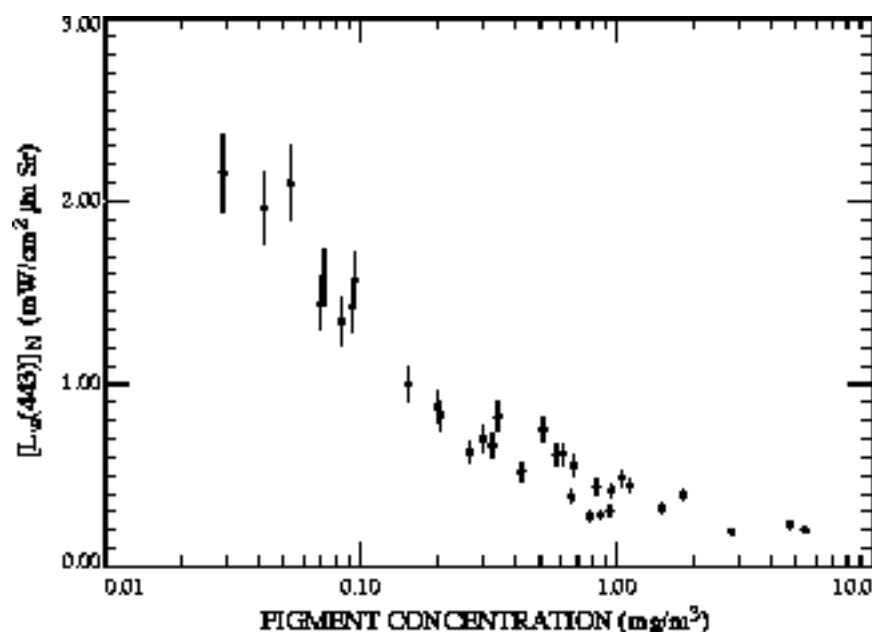


Figure 4a. Normalized water-leaving radiance at 440 nm; as a function of pigment concentration. Redrawn from Gordon *et. al.* [1988].

central role in our retrospective characterization. The most important property is the dependence of the normalized water-leaving radiance $[L_w]_N$ (Appendix 1) on the phytoplankton pigment concentration, C . The variation of $[L_w(i)]_N$ with C for Case 1 waters for CZCS Bands 1, 2, and 3, redrawn from Gordon *et. al.* [1988], is presented in Figures 4a, 4b, and 4c, respectively. We see that unlike Band 1, $[L_w(i)]_N$ for Bands 2 and 3 is only weakly dependent on C for $C > 0.3 \text{ mg} / \text{m}^3$ -- the clear water radiance behavior described in Appendix 1. These figures suggest the range of variation in $[L_w(i)]_N$ to be expected for Case 1 waters. The narrow range seen for Bands 2 and 3

allows examination of the short-term sensitivity fluctuations. Caution is required, however, in coccolithophore blooms where $[L_w(i)]_N$ can be significantly larger [Viollier and Sturm, 1984], e.g., at high concentrations of detached coccoliths $[L_w(i)]_N$ would be off the scale of Figures 4b and 4c. Such cases are termed "high backscattering" waters. Gordon *et al.* [1988] have developed a simple, semi-analytical model that explains the dependence of $[L_w(i)]_N$ on C , and attributes the "noisiness" of the relationship to natural variations in the scattering properties of phytoplankton.

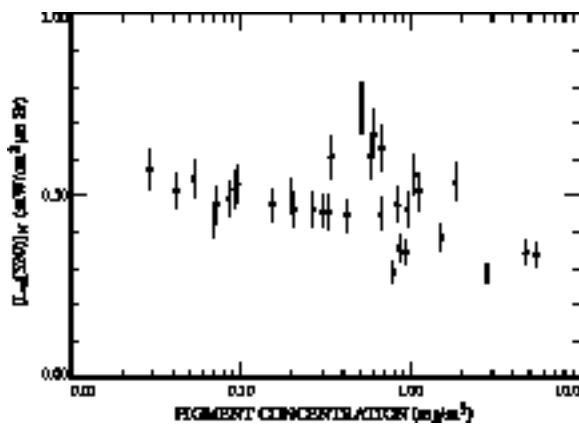


Figure 4b. Normalized water-leaving radiance at 520 nm as a function of pigment concentration. Redrawn from Gordon *et al.* [1988].

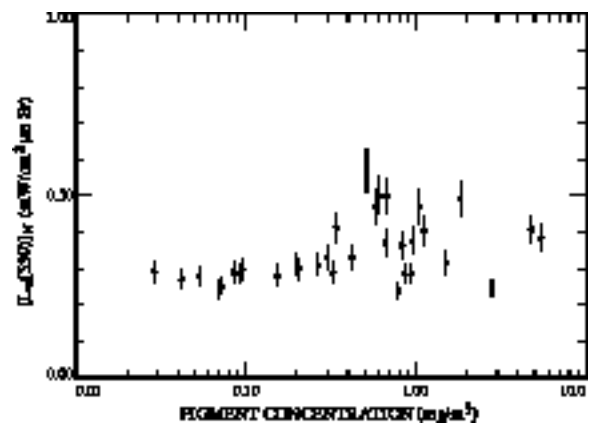


Figure 4c. Normalized water-leaving radiance at 550 nm as a function of pigment concentration. Redrawn from Gordon *et al.* [1988].

3.1.3.3 Basic Processing of the Global Data Set

Before beginning our retrospective examination of the CZCS radiometry, it is useful to review the procedures used in processing the CZCS global data set, since we use the same procedures in much of our analysis. The basic CZCS atmospheric correction algorithm is reviewed in Appendix 1. The quantities required for atmospheric correction of CZCS are given in Table 8. We shall discuss each in detail.

Table 8. Quantities and auxiliary data required for atmospheric correction.

Quantity	Auxiliary Data
$L_t(i)$	$S(i, G), I(i, G), k(i, G), g(i, t),$ Glint Mask, Cloud Mask
$L_r(i)$	$\langle F_0(i) \rangle, \tau_{O_z}(i), W, P_0$
$L_a(i)$	$L_w(4)$

Before processing can begin we need the calibrated radiances. These are found by applying the (nominal) calibration slopes and intercepts listed in Table 7. The results of the initial validation cruises suggested that the sensor calibration was in error by $\sim \pm 5\%$ depending on the band [Gordon et al., 1983, Gordon et al., 1983]. Thus the radiance computed via Eq. (4) must be corrected to account for this error. However, since the magnitude of the calibration error was assessed by applying the algorithms and comparing the results to sea truth, it is dependent on the details of the algorithm. In particular, it depends significantly on the computation of L_r ; usually the largest contributor to L_t . As described in Appendix 1, L_r was initially computed [Gordon et al., 1983] using the single scattering approximation; however, that procedure was later replaced by a computation based on an exact solution of the radiative transfer equation including polarization [Gordon, Brown and Evans, 1988]. L_r is also directly dependent on the extraterrestrial solar irradiance. The values of the constants $k(i, G)$ in Eq. (4) are provided in [Gordon, Brown and Evans, 1988] when the exact multiple scattering solution is used and $\langle F_0(i) \rangle$ is taken from Neckel and Labs [1984]. These were used in the initial stages of this work. The new values resulting from our present analysis are presented later in this paper (Table 10). The further modification of $L_t(i)$ required to account for the long-term loss in sensitivity of CZCS, i.e., $g(i, t)$ in Eq. (5), was also revised on the basis of the present analysis. In the initial stages of this work, we used the values proposed in Gordon et al., [1983].

Pixels contaminated by sun glint or clouds must be deleted from the imagery. A crude sun glint mask is provided by using the Cox and Munk [1954] model of surface roughness with a nominal windspeed of 6 m/s and rejecting pixels for which $L_g / F_0 > 0.005 \text{ Sr}^{-1}$. This removes most of the glint from pigment fields; however, some still remains in the water-leaving radiances near the edge of the mask. The cloud mask is made by a radiance threshold in an additional low-sensitivity spectral band at 750 nm. If $L_t(750 \text{ nm}) > 2 \text{ mW} / \text{cm}^2 \mu \text{mSr}$ the pixel is rejected as cloud contaminated. Note that no atmospheric correction is applied to $L_t(750 \text{ nm})$ and, in fact, no correction was made for the variation in cloud radiance with solar zenith angle or earth-sun distance. The increased saturation radiances for the SeaWiFS and MODIS bands allow the affects of clouds on adjacent pixels to be determined.

The aerosol radiance at 670 nm, $L_a(4)$, is determined by assigning $L_w(4) = 0$. This should introduce little error for low pigment concentrations; however, it will cause the retrieved $L_w(i)$ to be slightly underestimated for $C > 1 \text{ mg} / \text{m}^3$ since $L_w(4)$ will be

interpreted by the atmospheric correction algorithm as additional $L_a(4)$. Availability of the 750 nm channel for SeaWiFS and MODIS allows independent assessment of $L_w(670)$.

Finally, the quantities $\epsilon(1,4)$ can only be estimated over clear water [Gordon et al., 1983; Gordon and Clark, 1981]. Experience with the clear water procedure leads us to believe that with the exception of cases of continental haze extending out over the ocean from land, $\epsilon(2,4) \sim \epsilon(3,4) \sim 1$. Thus, we have chosen $\epsilon(2,4) = \epsilon(3,4) = 1$. $\epsilon(1,4)$ has been set equal to 0.95, i.e., 5% < 1, to account for multiple scattering as suggested by Gordon and Castaño [1987]. The assumption of fixed values for the ϵ factors is the *weakest link* in the global processing. This assumption will not be needed for SeaWiFS and MODIS since the addition of near IR channels allows the atmospheric correction to be determined given a wide range of chlorophyll concentration. The atmospheric correction yields the normalized water-leaving radiances $[L_w(i)]_N$ (Appendix 1) which are then inserted into the bio-optical algorithms [Gordon, 1990; Gordon et al., 1983] and the diffuse attenuation coefficient, $K(490)$, algorithm [Austin and Petzold, 1981] to derive fields of C and $K(490)$. Ken Carder's algorithms will be used to compute these parameters for SeaWiFS and MODIS.

A variation of the Mueller sensor overshoot algorithm [Mueller, 1988] was used to mask most of the erroneous computations that appear on the down scan (east) side of clouds from the slower-than-desirable recovery of the sensor after saturation. Each future sensor will have a specific strategy to either mask or possibly compensate for sensor overshoot. The initial approach for SeaWiFS will be to provide a mask to invalidate processing for pixels subject to sensor overshoot. Identification of these pixels will be made based on a determination of whether the magnitude of $L_t(i)$ exceeds a threshold for the particular combination of detector and gain selection. Selection of an appropriate strategy for MODIS must await test results for the flight instrument. Experience gained during on orbit operation potentially will allow partial correction assuming the magnitude of the overshoot can be calculated within acceptable error limits.

The results of these procedures are images of

$$[L_w(i)]_N (i = 1, n), L_a, C, K(490)$$

Time series of the images can be compared as well as time series of specific satellite derived-*in situ* parameters. will provide a basis to validate algorithms and sensor behavior.

3.1.3.4 Retrospective Characterization --- Methods

Previous approaches to CZCS characterization [Gordon *et al*, 1983, Mueller, 1985] relied on reference data sets compiled from limited spatial sampling of CZCS images due to lack of availability of the global archive. In our study the *entire* CZCS 4 km global archive was processed using the atmospheric correction procedures presented above. As each image is processed, C , $K(490)$, and $[L_w(i)]_N$ fields are combined with the data validity masks (clouds, glint, sensor tilt, sensor overshoot) to remove pixels containing contaminated samples. SeaWiFS and MODIS will also use global image sets spanning appropriate time ranges; this is anticipated to include increments of one year of data. Shorter time increments will be used following launch.

The remaining pixels are still subject to radiance contributions from high backscattering regions (e.g., coccolithophore blooms, Case 2 waters, etc.), atmospheric correction algorithm limitations at high latitudes, or the possibly erroneous assumption the $\varepsilon(1,4)$ is always unity. To minimize effects of nonentity $\varepsilon(1,4)$, $L_a(4)$ is required to be $<1.0 \text{ mW} / \text{cm}^2 \mu\text{mSr}$. The effects of the others, with the exception of high backscattering areas, are minimized by stratifying the data as they are mapped onto a global grid. For the present effort a global, equal-angle grid (with a constant latitude increment and a constant longitude increment) with 2048 columns and 1024 rows is defined resulting in a pixel $\sim 18 \text{ km}$ on a side at the equator. SeaWiFS and MODIS will use an 9km, 4096 x 2048 (4km when computer peripheral capacity permits) equal area grid (see Appendix 2). Another substantive difference between CZCS and SeaWiFS, MODIS is the definition of the *data day*. (see Appendix 3). The CZCS data day is based on data acquired within a 24 hour period; the present approach includes data within specific meridians, e.g. (180 - 180). This approach also eliminates temporal ambiguities within the daily field as well as provides a simple method to relocate the boundary, e.g. (0-0 or any other arbitrary meridian) without reprocessing the daily fields). Masks are prepared at the same scale defining a latitude range within 50° of the equator and for water depths exceeding 2 km. These masks address latitude correction limits and restrict ocean retrievals for the most part to Case 1 waters. Presence of high backscattering areas introduced by coccolithophore blooms is not addressed. Finally only scenes with sensor $G=1$ are included. These accounted for 64% of all of the CZCS imagery. Several investigators, e.g. Balch, have proposed algorithms for identifying and quantifying coccolithophore blooms. Thus next generation algorithms will be able to remove these areas from consideration as targets for sensor calibration.

$[L_w(i)]_N$ for pixels meeting these criteria are combined onto the global map and binned in increments of $0.01 \text{ mW} / \text{cm}^2 \mu\text{mSr}$. This permits daily fields to be combined, a capability that increases the extent of ocean included in a given sample. Increments of 10 days are used in this study. $[L_w(i)]_N$ for each 10 day increment is

then binned in $0.01 \text{ mW} / \text{cm}^2 \mu\text{mSr}$ cells and the resulting histograms are collected into time series to determine temporal trends. Selection of the 10 day collection interval for CZCS is based on the relatively small global fraction recorded by the sensor on a given day. Intervals selected for SeaWiFS and MODIS will range between 2 and 8 days where the selection will be governed by availability of global cloud free observations.. The general approach that we have taken to characterize the CZCS is to apply the algorithms to the whole data set and force the results to satisfy certain constraints. For example, considering the way we have stratified the data, the frequency distributions of $[L_w(2)]_N$ and $[L_w(3)]_N$ should have a peak near the clear water values of ~ 0.50 and $0.30 \text{ mW} / \text{cm}^2 \mu\text{mSr}$ for Bands 2 and 3, respectively (Figures 4b and 4c). The values of $g(i, t)$ and $k(i, G)$ in Eq. (5) are chosen to *force* this to happen. Since we are using a global, multi-year dataset, this procedure is equivalent to the assumption that there is *no interannual variability* in $[L_w(2)]_N$ and $[L_w(3)]_N$ on a *global scale*. Availability of the time series derived for the *in situ* observations will minimize this limitation for SeaWiFS and MODIS.

As described above, after the algorithms are applied to each ten-day period, histograms of $[L_w(2)]_N$ and $[L_w(3)]_N$ are produced; examples of such histograms are provided in Figures 5a - 5d. These samples were from near the *end* of the characterization process described below, i.e., after reasonably accurate values of $k(i, G)$ and $g(i, t)$ had already been obtained. These figures have been prepared by further stratifying the data to include only pixels with $L_a(4) < 0.3$ or $1.0 \text{ mW} / \text{cm}^2 \mu\text{mSr}$. For $L_a(4) < 0.3 \text{ mW} / \text{cm}^2 \mu\text{mSr}$ in Figures 5a and 5b the maximum frequency of occurrence of $[L_w(2)]_N$ and $[L_w(3)]_N$ is seen to be very near the clear water values. When the maximum allowable $L_a(4)$ is increased to $1.0 \text{ mW} / \text{cm}^2 \mu\text{mSr}$, the distributions are seen to broaden a little but the maximum in the distribution is unchanged. These observations are characteristic of the global data prior to about August 1981 and are consistent with our assumption that there is little interannual variability in $[L_w(2)]_N$ and $[L_w(3)]_N$ on a global scale.

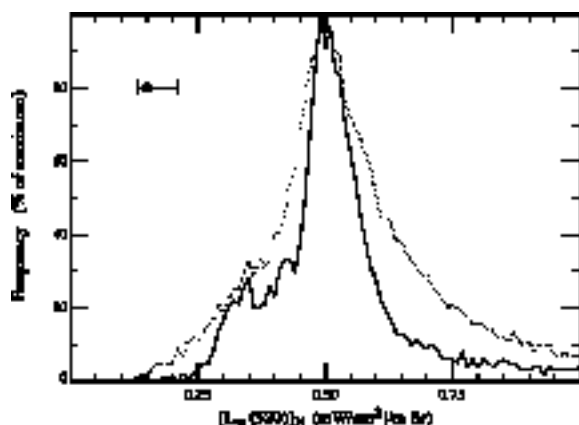


Figure 5a. Global frequency distribution of normalized water-leaving radiance at 520 nm for a 10-day period in mid June 1981. Horizontal error bar represents the estimated error around the computed value (dot). Solid and dotted lines are for $L_a(4) < 0.3$ and $1.0 \text{ mW} / \text{cm}^2 \mu\text{mSr}$, respectively.

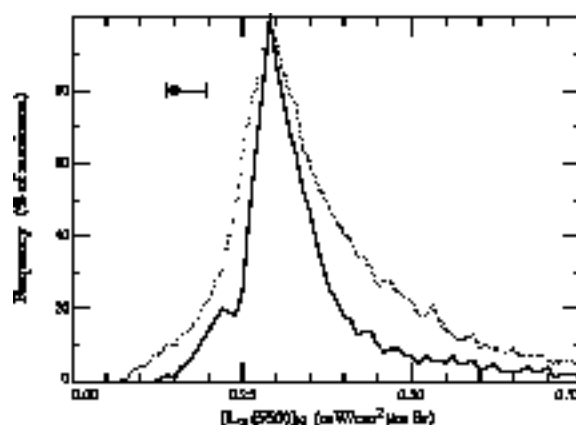


Figure 5b. Global frequency distribution of normalized water-leaving radiance at 550 nm for a 10-day period in mid June 1981. Horizontal error bar represents the estimated error around the computed value (dot). Solid and dotted lines are for $L_a(4) < 0.3$ and $1.0 \text{ mW} / \text{cm}^2 \mu\text{mSr}$, respectively.

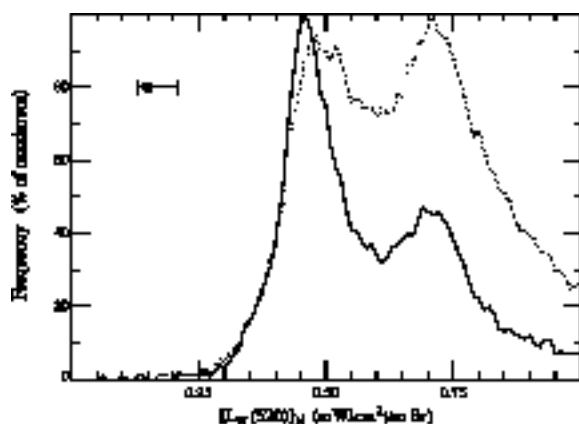


Figure 5c. Global frequency distribution of normalized water-leaving radiance at 520 nm for a 10-day period in early October 1981. Horizontal error bar represents the estimated error around the computed value (dot). Solid and dotted lines are for $L_a(4) < 0.3$ and $1.0 \text{ mW} / \text{cm}^2 \mu\text{mSr}$, respectively.

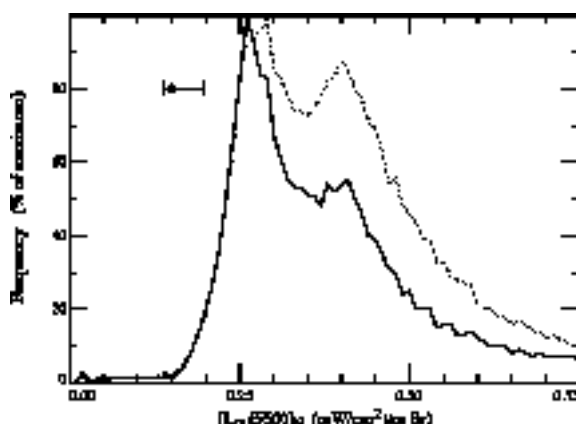


Figure 5d. Global frequency distribution of normalized water-leaving radiance at 550 nm for a 10-day period in early October 1981. Horizontal error bar represents the estimated error around the computed value (dot). Solid and dotted lines are for $L_a(4) < 0.3$ and $1.0 \text{ mW} / \text{cm}^2 \mu\text{mSr}$, respectively.

After August 1981, two distinct maxima are often observed in the histograms as seen in Figures 5c and 5d. In Band 2 the double peak was observed 30-50% of the time after August 1981, while for Band 3 the occurrence was about half as frequent. The occurrence of the second maximum far above the expected clear water value causes a small shift (~ 1 DC) in the clear water maximum. This second maximum is reduced if the data are limited according to $L_a(4) < 0.3 \text{ mW} / \text{cm}^2 \mu\text{mSr}$; however, it does not disappear. Since this second maximum was not seen in the processing for the first 2.5 years, we conclude that it is instrumental, i.e., an artifact of the instrument and *not* in the actual water-leaving radiances. It is only possible to make such a claim because the data being examined are global and multi-year.

3.1.4 Variance or uncertainty estimates

3.1.4.1 Observation frequency for useful *In situ* - satellite matchups

Application of the match-up database focuses on changes in sensor performance at time scales longer than six months. The expected scatter in the satellite-*in situ* residuals will be large with respect to changes that will likely occur during normal sensor operations. A number of samples spaced over an extended time period is needed to establish the temporal trends. Table 9 shows the fraction of useful retrievals available for two buoys extracted from the NOAA-9 MDB. If the useful fraction of 25% is indicative of the fraction expected for the MODIS timeframe, then the frequency of MODIS observation of a given buoy location, 1 every 2 days, will yield approximately 90 samples/year. These will be distributed in scan angle by the orbit repeat frequency, every 4 - 8 days depending on magnitude of scan angle yielding 10 - 20 sample/year for a given buoy for a range of scan angles. Data from a number of buoys is required to increase data density and to develop error statistics given the uncertainties present in both the *in situ* and satellite observations. Where multiple buoys are available, the satellite becomes a "transfer standard" permitting relative calibration checks between *in situ* platforms. This approach has been used to check buoy calibrations across retrieval and re-deployment events and to determine when drifting buoys have exceeded their useful lifetime.

Table 9. Buoy retrieval statistics for fixed mooring and NOAA-9 AVHRR

Time	Total Observations over 4 years	Cloud-free Observations	% Useful Observations
Buoy 1-all	1731	390	23%
day	868	191	22%
night	863	199	23%
Buoy 2-all	1686	526	29%
day	955	250	26%
night	831	276	33%

The smaller daytime percentage reflects a loss of coverage due to sunglint. The AVHRR sensor collects full earth coverage within a 24 hour period. Even with this extensive coverage, approximately 75% of the *in-situ* observations do not contribute useful matchups with the satellite due to clouds and sun glint. Matchup retrievals obtained from MODIS will have a lower percentage of useful observations due to the restricted scan coverage where 2 days are required to observe the entire surface of the earth. Thus visible wavelength matchup database will be sparsely populated. These data together with the more extensive data sets obtained during cruise periods form the test data sets that will be used to quantify algorithm/sensor performance.

Other investigators (e.g. Abbott, Lewis) have discussed the possibility of deploying optical drifters. Data provided by this type of instrument have the potential of greatly increasing the number of *in-situ* observations available to characterize the MODIS sensor and algorithm performance. While these observations will not provide the same level of calibration and continuity as the fixed buoy instruments, they provide a greater range of geographical and geophysical observations than the limited set of fixed buoy observations. In a relative sense, changes in retrieved radiance can be tested against parameters such as scan angle, time, location... Availability of a large number of *in situ* sources provides a dispersed set of reference information that permits changes in retrieved radiance to be traced to the sensor, algorithm, or *in-situ* location.

Brown in his ATBD discusses the difference between skin and bulk SST. Skin temperature is measured using a downward looking radiometer deployed on ships or buoys. While there are likely to be more IR than visible instruments, the overall

situation of a limited number of locations observed by radiance instruments will be true for IR observations as well as visible. Data describing the larger geophysical setting for the most part is provided by fixed or drifting buoys measuring bulk temperature with a thermistor mounted at a depth of one to several meters. This larger data set again provides a context check for the limited number of more accurate radiometric measurements.

3.1.4.2 Error sources

Error calculations for the MODIS ocean color atmospheric correction approach are presented in H. Gordon's ATBD and will not be reproduced here. Application of the matchup database will be directed towards trying to differentiate between errors due to the atmospheric correction process, sensor characterization and *in situ* measurements. We will utilize comparisons between expected cross scan behavior for the 500 nm bands and observed change to indicate change in mirror reflectivity or polarization effects. Figures are presented later showing behaviour of the long wavelength AVHRR IR bands as a function of scan angle and latitude. Comparisons of similar fields obtained from SeaWiFS, OCTS, AVHRR and MODIS will be used in conjunction with ancillary data to identify locations and conditions where sensor and algorithm performance is satisfactory. Where not, the data sets will help identify potential sources of error. The remainder of this section discusses how errors were constrained for the retrospective CZCS characterization.

It is useful at this point to estimate the error in histograms such as those presented in Figure 5 resulting from errors in the atmospheric correction. The correction errors can come from two sources: error in the assumption that $L_w(4) = 0$; and error from the assumption that $\varepsilon(2, 4) = \varepsilon(3, 4) = 1$. The former is not considered important here because the data are stratified in such a way as to avoid the high pigment concentrations where this assumption breaks down, i.e., coastal areas. The latter approximation is more likely to be incorrect and we shall assume that it is the sole source of the error.

If $L_a(4)$ is sufficiently small, we are justified in assuming single scattering for the aerosol component. In particular, this is a very good approximation when $La < 0.3$

$mW / cm^2 \mu mSr$. Since the error in $L_w(i)$ is assumed to be dominated by the error in the assumed value of $\varepsilon(1,4)$, we have from Appendix 1

$$(tL_w(i)) = \varepsilon(i,4) L_a(4) \frac{F'_0(i)}{F'_0(4)} \quad (6)$$

The results were

$$\varepsilon(2,4) = 1.0 \frac{+0.10}{-0.05}$$

$$\varepsilon(2,4) = 1.0 \frac{+0.10}{-0.05}$$

From $(tL_w(i))$ we must compute $[L_w(i)]_N$, the error in the normalized water-leaving radiance. Using the definition of $[L_w(i)]_N$,

$$[L_w(i)]_N = \frac{t L_w(i)}{\cos \theta_0} \exp \left(\frac{\tau_r(i)}{2} + \tau_{oz}(i) \right) \frac{1}{\cos \theta_0} + \frac{1}{\cos \theta_v} \quad (7)$$

Taking the mean values of $\cos \theta_0$ and $\cos \theta_v$ to be 0.75, we find

$$[L_w(i)]_N = 1.65 t L_w(i)$$

$tL_w(i)$ from Eq(6) and the limits determined for $\varepsilon(1,4)$, with $L_a(4) < 0.3 mW / cm^2 \mu mSr$, yield

$$[L_w(2)]_N = \frac{+0.059}{-0.029}$$

$$[L_w(3)]_N = \frac{+0.048}{-0.011}$$

in $mW / cm^2 \mu mSr$. These are the error bars shown on Figures 5a to 5d. For larger $L_a(4)$, $[L_w(2)]_N$ will be proportionally larger. In the case of nonmaritime aerosols, $\varepsilon(1,4)$ can be much larger than 1, and when they occur the error can be significantly larger. Limiting the geographical location of the imagery used, as we did, should minimize their occurrence. Note, however, that since an underestimation of $\varepsilon(i,4)$ will cause the retrieved $L_w(i)$ to be too large, these nonmaritime aerosols will *only* affect the *upper tail* of the $[L_w(i)]_N$ distribution.

It is comforting to note that the "halfwidth" of the $[L_w(i)]_N$ distribution is approximately equal to the estimated uncertainty in the analysis resulting from the assumptions regarding constancy of $\varepsilon(1,4)$. Also, when the $L_a(4)$ limit is increased to $1.0 mW / cm^2 \mu mSr$, Figures 5a and 5b show that the "width" of the distribution does not increase significantly. This suggests that our estimate of $[L_w(i)]_N$ in this case (three times the error bars shown on Figure 5) is very conservative. Comparing the cases for the two $L_a(4)$ limits in Figure 5 shows that the positions of the $[L_w(i)]_N$ maxima change very little with the $L_a(4)$ limit. This is significant because using the $L_a(4) < 1.0 mW / cm^2 \mu mSr$ limit allows a considerable increase in the amount, and therefore the representativeness, of the data that is used in our examination.

The possibility of radiance contributions from whitecaps on the sea surface has been ignored. Their contribution will be interpreted by the atmospheric correction algorithm as additional aerosol, i.e., $L_a(4)$. Since whitecaps are white, their effect will be to move the actual ε -values toward 1, i.e., toward the value we have assumed. Thus, although whitecaps will in general interfere with the correction algorithm, in the present application they will not because of our ε -assumption. Algorithms for SeaWiFS and MODIS include corrections for whitecaps.

3.1.4.3 Initialization and Long-Term Degradation

The CZCS sensitivity seemed to vary over a range of time scales (days to years); however, since the variability over long time scales (months to years) appeared to be much larger than that over short time scales (days to weeks) we chose to separate the long-term component from the total variation. The methodology for such a separation follows.

The global data set is processed assuming that $\varepsilon = 1$ and accounts for multiple scattering by lowering $\varepsilon(1,4)$ by 5%. Gordon's SeaWiFS/MODIS atmospheric correction algorithm directly calculates ε values based on use of the near IR bands; all other bands are calibrated relative to these atmospheric bands. Thus it is important that atmospheric measurements are made to provide reference information to calibrate the near IR bands. Initially, it is assumed that there is no degradation ($g(1,t) = 1$) and that $k(i,G) = 1$ in Eq. (5). Thus the first task is to determine approximate values for these quantities. This is effected in several steps.

Step 1: the algorithms are applied to the global data set to form the ten-day histograms of $[L_w(2)]_N$ and $[L_w(3)]_N$.

Step 2: $g(i,t)$ is adjusted to force agreement between the radiance with maximum frequency and the corresponding clear water values of $[L_w(2)]_N$ and $[L_w(3)]_N$. In situations such as that shown on Figure 5c in which the second peak in the distribution of $[L_w(2)]_N$ occurs more frequently than the peak near the clear water value, one must decide which peak to choose to force the agreement. This was done by objectively looking at the time series of histograms and choosing the peak that produced smaller discontinuous jumps in the calibration. Typically, this was the smaller peak, but not always.

Step 3: $g(i,t)$ is smoothed to retain only the long-term component for Bands 2 and 3.

Step 4: $g(i,t)$ is used to reprocess the entire data set again. If the retrieved $[L_w(2)]_N$ and $[L_w(3)]_N$ histograms do not have their maximum at the clear water values, there is inconsistency between the $k(i,G)$, $g(i,t)$, $\varepsilon(1,4)$, $[L_w(2)]_N$ and $[L_w(3)]_N$. We then adjust $k(i,G)$, reprocess, and continue as many times as required to achieve the desired consistency. At this point, we assume that the "correct" values of $k(i,G)$ and $g(i,t)$ for Bands 2 and 3 have been obtained.

Step 5: To obtain $k(i,G)$ and $g(i,t)$ for Band 1, we applied the atmospheric correction algorithm and the clear water-radiance method described in Appendix 1 to obtain $\varepsilon(1,4)$ and therefore $[L_w(1)]_N$ for imagery obtained simultaneously with the NET post launch validation cruises (nine stations). $k(i,G)$ and $g(1,t)$ are readjusted to force agreement between CZCS-retrieved and ship-measured values of $[L_w(1)]_N$.

A brief review of the specific steps taken to compute $k(1,1)$ illustrates the uncertainty introduced when an incomplete set of measurements is available to characterize the atmospheric and in-water radiances. No atmospheric measurements were available to assist in the determination of $\varepsilon(1,4)$. Thus we rely on Steps 1 - 4 to determine $k(2,1)$ and $k(3,1)$. We then require that ship-observed radiance data be available adjacent to a source of clear water. This source was a warm core ring for the June 1979 images of the Middle Atlantic Bight. The $\varepsilon(1,4)$ are computed (Appendix 1) in the ring and transferred to the station location assuming no spatial variation. Satellite $[L_w(1)]_N$ is computed and the $k(1,1)$ adjusted until a match with the ship observed radiance is obtained. A consistency test is available; since the $\varepsilon(i,4)$ are not computed at the station, the satellite- and ship-observed $[L_w(2)]_N$ and $[L_w(3)]_N$ must agree within the expected error. A source of uncertainty is introduced due to the station locations. We now believe that they were positioned in a transition region between maritime, $\varepsilon(3,4) \sim 1$, and continental, $\varepsilon(3,4) = 1.17$, aerosol distributions. In *Gordon et al. [1983]*, the entire Middle Atlantic Bight was assumed to be dominated by the continental aerosol. In this work, where the $k(i,1)$ for Bands 2 and 3 are fixed by the global data set, computation of $\varepsilon(1,4)$ in the clear water ring area produced an ε midway between the maritime and continental aerosol types. The difference in $\varepsilon(1,4)$ values is a principal difference in the $k(i,1)$'s calculated for these two calibration efforts. Without an accompanying atmospheric data set, there is no method to resolve the ambiguity in $\varepsilon(1,4)$ and therefore in $k(i,1)$. [Recall that the global data set is processed assuming a maritime aerosol]

Step 6: To extend $g(1,t)$ beyond June 1979, we assume that the $g(i,t)$ vary temporally in a piece-wise linear manner and use ship track pigment data, along with the atmospheric correction algorithm, to force the CZCS-derived pigment to agree with the ship-measured pigment.

The previous six steps yield a detailed characterization of the state of the CZCS in Bands 2 and 3 relative to Band 4, i.e., Band 4 was initially assumed to have $g(4,t) = 1$ and $k(4,1) = 1$. For Band 1 the characterization is coarse because of the paucity of surface data.

Step 7: Once the $k(i,G)$ for $G=1$ are established, the remainder of the k 's are set by comparing the $L_t(i)$ for images that contain changes in G . Once suitable images are located, the adjacent scan lines with different values of G are edited to remove clouds and the counts on each scan line averaged. The count ratio corresponds to the k -ratio for the two gains. Where scan lines include a range of counts, both the $S(i,G)$ and $I(i,G)$ can be adjusted. This latter approach is used to make the small adjustments

seen in $I(i, G)$, yielding a slightly different set of I -values (the $I(i, G)^*$ values given in Table 7) from the original values determined during pre-launch tests [Ball Aerospace Division, Boulder CO, 1979]. The gain ratios are considered to be temporally invariant. Limitations with this assumption are discussed in a later section. The resulting values of $k(i, G)$ are presented in Table 10. Importantly, this procedure is applicable in normalizing gains for the various detectors for any given band in the SeaWiFS and MODIS instruments.

Table 10. $k(i, G)$ determined as described in the text

Band	Gain			
	1	2	3	4
1	1.018	1.021	1.011	1.020
2	0.982	0.983	0.988	0.972
3	0.974	0.963	0.947	0.950
4	1.008	1.020	1.016	1.010

3.1.5 Retrospective Characterization -- CZCS Results

The Nimbus-7 CZCS sensor at present provides the only example of temporal behavior for an ocean color sensor. Many of the assumptions that were made for this sensor will not be required for MODIS. For example, an important lesson learned from the CZCS experience is the necessity of making periodic, high quality *in situ* observations of both the ocean and atmosphere. The MOBY and MOCE buoy and ship observations will provide the necessary *in situ* observations.

3.1.5.1 Long-term Variation

The long-term degradation resulting from the CZCS studies is shown in Figure 6. We now discuss the individual bands in detail.

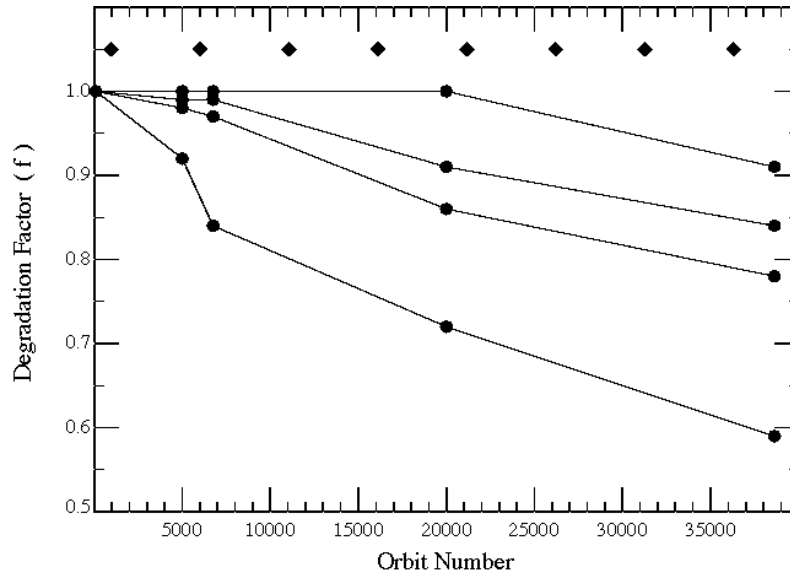


Figure 6. Long-term degradation of the sensitivity $f=g^{-1}$ of the 4 CZCS bands. Wavelength increases from 443 nm to 670 nm from bottom to top. The diamonds correspond to the beginning of the calendar year starting with 1979.

3.1.5.2 Band 4

There are no data sets available to use as a reference for Band 4, so a detailed examination was not possible. This band was initially assumed to be stable, i.e., the degradation was sufficiently small to be ignored. The aerosol component, $L_t(4) - L_r(4) = L_a(4)$, was compared for early and late in the mission for scenes with very clear atmospheres. After launch the lowest observed $L_a(4)$ radiances were order of $0.20 \text{ mW} / \text{cm}^2 \mu\text{mSr}$ at a time when the stratospheric atmospheric aerosol burden was very low [Hofmann, 1988]. Subsequent volcanic activity, e.g., Mt. St. Helens and El Chich'on, raised the level of aerosols, and $L_a(4)$ remained at or above the level observed during 1979, which represents a conservative floor. When data from 1985-1986 period were analyzed, $L_a(4)$ was found to be reduced to near zero or even have negative values. Because of this, a small degradation is introduced into Band 4 beginning in late 1982 (Orbit 20000) to raise the minimum retrieved $L_a(4)$ in the 1985-1986 time period to correspond to the values observed for 1979. The choice of Orbit 20000 is arbitrary but precedes the various strategies, e.g., sensor heating, that were employed to keep the CZCS operational through mid 1986. It will be important to acquire ship based atmospheric measurements for SeaWiFS and MODIS to minimize the impact on application of the atmospheric correction algorithm due to lack of

specificity on the aerosol properties present during sensor initialization and post-launch calibration.

3.1.5.3 Bands 2 and 3

The degradation rates for the two green bands are relatively constant over the sensor life. The overall character of the degradation suggests the loss of sensitivity increases with decreasing wavelength.

3.1.5.4 Band 1

Band 1 experienced the largest change in sensitivity losing approximately 40% by mid 1986. It experienced a higher initial degradation rate than the other bands. This initial rate then accelerated during the period from Orbit 5000, late October, 1979, to Orbit 6750, late February, 1980. This change in degradation behavior was noted during processing of the global data set when unrealistically high chlorophyll values were computed. Histograms of $[L_w(1)]_N$ were computed and the $g(1, t)$ adjusted such that the overall position of the histograms remained invariant through the time period. This period corresponded to the time when efforts were made to recover use of the CZCS thermal channel by heating the cooler in an attempt to remove suspected contamination. This 4+ month period accounted for a 12% sensitivity loss or nearly 1/3 of the total sensitivity loss over the sensor lifetime. Availability of the global record allowed this time interval to be characterized and this represents a significant point of departure from results of previous degradation studies. By March 1980, the degradation rate for Band 1 decreased and remained approximately uniform for the rest of the mission.

3.1.5.4 Short-term Fluctuations

As described earlier, fluctuations with much shorter time scales are superimposed on the long-term variation. Short-term fluctuations for Bands 2 and 3 are shown in Figure 7, which shows $[L_w(i)]_N$, the difference between the derived $[L_w(i)]_N$ and the clear water values. The short-term fluctuations in Bands 2 and 3 suggest that the sensor was stable at the ± 1 DC level through the Fall of 1981 (81210, Orbit 14000). The 10-day histograms are essentially invariant from launch to this date and exhibit a single peak with a rapidly increasing initial shoulder and a broader trailing tail that most likely represents aerosol distributions with $\varepsilon(i, 4) > 1$. All atmospheric corrections assumed that $\varepsilon(i, 4) = 1$ so any locations containing aerosols with larger ε 's will be interpreted as having larger than nominal $[L_w(i)]_N$ retrievals. In our analysis, the peak of the

histogram must be positioned at the clear water value when the sensor calibration is correctly set. Figure 7 shows that we actually set the calibration for these two bands about 1 DC too low. After the Fall of 1981, large excursions are seen in both the Band 2 and 3 curves. Individual 10 day histograms tend to show double peaks (Figures 5c and 5d) suggesting that the sensor has two potential states for a given channel (e.g., values of $S(i, G)$ or $I(i, G)$) that are possible at instrument activation. While the fluctuations are similar, Figure 7 shows that the timing and magnitude of the changes differ for the two channels. The CZCS was turned off between sampling sessions to conserve power. As such, each time the instrument was activated, the electronics established a new zero reference. We suspect that this zero reference was not stable after Fall 1981 and that this is the root of the double maxima in the $[L_w(i)]_N$ histograms.

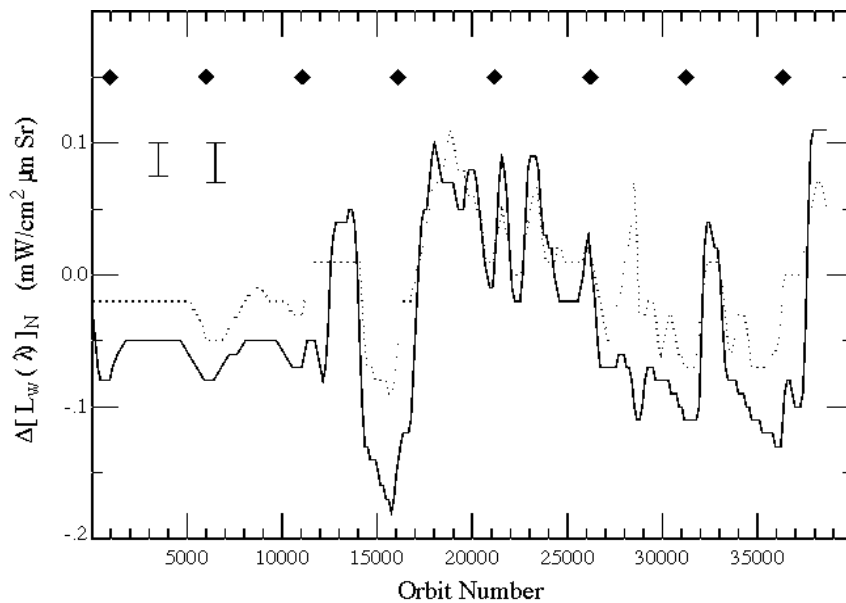


Figure 7. Calibration variation in the bands at 520nm (solid line) and 550 nm (dotted line). The diamonds correspond to the beginning of the calendar year starting with 1979. The two "error bars" at the upper left are the radiances corresponding to one digital count of the CZCS (shorter bar is for 550 nm)

Since it is possible for events not related to the sensor to introduce histogram shape changes, checks were conducted to test likely possibilities. Atmospheric correction assumed that $\varepsilon(i, 4)=1$ could be globally applied. Any error in this assumption would be interpreted as enhanced or depressed $[L_w(i)]_N$; however, Figure 5 suggests that an error in this ε assumption would not materially shift the *peak* of the $[L_w(i)]_N$ distributions by more than $\sim 1DC$, indicating that the selection criteria lead to a robust

estimate of most frequently occurring value of $[L_w(i)]_N$. In addition, for situations where there are multiple maxima, their separation in Bands 2 and 3 is difficult to explain by non-unity ϵ 's. For example, the separation between the two peaks in Band 2 (Figure 5c) is 40% greater than their separation in Band 3 (Figure 5d). This would require an unphysically large increase in $\epsilon(2,4)$ over $\epsilon(3,4)$. Another check examined whether the large increase in the short-term variation that occurred in early 1982 could be related to the eruption of El Chich'on or some other aerosol source. An additional mask was used to remove a zonal band from 15° N to 32° N, an area principally affected by the El Chich'on aerosols. No changes were noted in the standard global histograms compared to the zonally masked set. These tests suggest that likely external sources of histogram shape perturbation resulting from increased atmospheric turbidity or non-unity ϵ 's do not have a significant impact on the histogram peak location. The procedure appears to be robust, allowing use of the clear water approximation as a surrogate calibration reference for the two CZCS green bands and as a stability reference for the sensor.

3.1.6 Concluding Remarks

In this section a retrospective look at the CZCS system-- sensor plus algorithms -- calibration is presented together with modifications appropriate to the upcoming ocean color instruments. This is based on (1) an examination of the few instances in which contemporaneous satellite data and surface measurements of water-leaving radiances were available, (2) an examination of the few instances in which contemporaneous satellite data and surface measurements of pigment concentration were available, and (3) processing of the entire CZCS data set invoking the hypothesis that there is no interannual variability in the global distribution of the normalized water-leaving radiances in Bands 2 and 3 and assuming that the radiometric sensitivity of Band 4 varied minimally during the entire mission. In the absence of atmospheric measurements, additional assumptions were required regarding the assumed nature of the aerosol over the ocean.

The results show that the calibration of Bands 2 and 3 displayed short-term fluctuations superimposed on a long-term decrease in radiometric sensitivity. The short-term fluctuations are highly correlated in Bands 2 and 3, but not completely correlated. It is remarkable, however, that when corrected for long-term variation, CZCS Bands 2 and 3 were stable to within $\pm DC$ for the first 2.5 years of the mission. The extrema of the Band 2 and Band 3 variations is approximately $\pm 2\%$ of the typical radiance (L_t)

observed for these bands. This error introduces an uncertainty of approximately a factor of 2 in the retrieved pigment concentration. The next generation of ocean color sensors, i.e., SeaWiFS and MODIS, are required to be calibrated on the ground with an absolute uncertainty of 3-5%. Clearly post launch vicarious calibration will be required to reduce the calibration uncertainty and provide acceptable uncertainties in the water-leaving radiances and the pigment concentrations.

The long-term variation is observed to decrease in magnitude with increasing wavelength. It is not seen in the calibration lamp data suggesting that it results from contamination on the optical surfaces directly exposed to the space environment. This suggests that future visible radiometers such as SeaWiFS should acquire frequent solar and lunar calibration data to ascertain sensor stability.

This analysis is based on processing the entire global data set twelve times to compute and validate the required coefficients.

Several lessons have been learned from the CZCS experience for application to future missions, e.g., SeaWiFS and MODIS. First, ocean color sensors should be designed with provisions to (1) insure the instrument is as radiometrically stable as possible, and (2) to provide on-board systems for monitoring the radiometric stability of the sensor which utilize the entire optical train. Next, an initialization process is required which provides atmospheric as well as water-leaving radiance measurements to accurately set and tune the system - sensor plus algorithms. Finally, at least one ocean observation station should be established in relatively clear water (void of strong horizontal variations in C that could lead to sub-pixel structure) to obtain daily measurements of $[L_w(i)]_N$ simultaneously with satellite overpasses. If such an ocean observation station had been in place during the CZCS era, the data would have yielded considerably more accurate global pigment fields.

3.2 Practical Considerations

3.2.1 Overview

Processing time for the ocean color algorithms is dominated by the atmospheric correction step. Experience using the SeaWiFS test data set has provided an opportunity to evaluate potential mechanisms to minimize processing time. One method that has been incorporated is to retain knowledge of the aerosol models chosen to be appropriate at the previous pixel and assume that the same models work at the current pixel. This assumption is tested, and if found to be true, results in a considerable savings since only two rather than twelve aerosols need to be evaluated.

Programs have been developed and submitted (Version 1) to the MODIS team (SDST) for compliance testing and delivery to the GSFC DAAC. Programs have been coded using C and Fortran 90 using EOS toolkit support routines and HDF-EOS file structures.

There are two Level-2 routines, one for SST and another for ocean color. Routines have been integrated from Brown for SST, Gordon for atmospheric correction, Carder for chlorophyll and other Case II water products, Abbott for fluorescence line height and chlorophyll fluorescence efficiency, Hoge for PUB and PEB, Clark for various Case I water properties. Binning programs have been developed to generate Level-3 equal area files for each of the products. These files can be binned into desired space and time resolutions required to meet both EOS standard and investigator specific resolutions. Finally a mapping program is available to produce gridded products.

3.2.2 Programming/Procedural Considerations

3.2.2.1 Overview

Processing estimates are currently based on Version 1 ocean color and SST PGEs executing on SGI R10000 200Mhz processors using IRIX 6.2, EOS toolkit calls and HDF-EOS file structures and MODIS test Level-1b input data. Granule processing times are based on a granule that contains no land and no clouds. The Version 1 programs have integrated the latest versions of the atmospheric correction and product algorithms supplied by the MODIS ocean team investigators.

3.2.2.2 SST and ocean color processing requirements

Processing times for SST and ocean color Level-2 and Level-3 programs are presented in Table 11. Individual granule processing times are presented for each step of the Level-2 and Level-3 processing path. Extrapolation to cpu estimates for processing an entire day's worth of data do not include savings due to skipping land or cloud contaminated pixels. It is expected that not processing these pixels will save upward to 2/3 of the total estimated time. These savings are not included since it is anticipated that increases in computational load that will occur once Version 2 algorithms are delivered and integrated will consume any savings. Ocean color times do not include time savings that are anticipated due to pixel to pixel correlation of the aerosol models; the global time estimated does include processing only daylight pixels.

Processing estimates forecast in the Version 0 ATBD allowed 5 100 megaflop processors and 12 100 megaflop processors respectively for the SST and ocean color algorithms. The present estimates are consistent with the original forecast since the 288 processor hours required for Level-2 ocean color can be delivered by 12 processors.

Table 11. Processing loads for SST and ocean color

process ID	Input File IDs ocncolor*=lw,dr1 ,dr2, qc ocncolor*_= name for each of the 36 ocean color products sst_**=d or n	Output File IDs ocncolor*=lw,d r1,dr2, qc ocncolor*_= name for each of the 36 ocean color products sst_**=d or n	maximum minutes of CPU processin g time per execution	No. of Exec. /day	MB of output created by processes active every day	maximum minutes of CPU to process a day of data
MODIS Level 2 Production Ocean Color Parameters	MOD02_L1B	ocncolor*	120.0	144.00	140400.0	17280
	ANC_ESI					
	DAS_SLU					
	DAS_SLV					
	DAS_WINDS					
	DAS_PSFC					
	DAS_SLP					
	DAS_PRSE					
	DAS_PRS_SPHU					
	DAS_PRS_TMPU					
	ANC_NMC_SURF					
	MOD03_L1A					
	MOD35_I2					
	MOD04					

process ID	Input File IDs ocncolor*=lw,dr1 ,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	Output File IDs ocncolor*=lw,d r1,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	maximum minutes of CPU processin g time per execution	No. of Exec. /day	MB of output created by processes active every day	maximum minutes of CPU to process a day of data
	MOD18_L3_REF_WK					
MODIS Ocean Color Space Binning	ocncolor*	ocncolr*_int_p	12.0	144.00	40089.6	1728
MODIS Ocean Color Production of Single Orbit	ocncolr*_int_p	ocncolor*_int_orb	0.1	432.00	18792.0	57.6
MODIS Ocean Color Production of Daily Composite	ocncolor*_int_orb	ocncolor*_int_dy	4.0	36.00	18792.0	144
Form weekly composite and reference field	ocncolor*_int_dy	ocncolor*_int_wk	3.3	5.14	NA	120
	ocncolor*_int_wk	ocncolor*_int_3w k	103.3	5.14	NA	3720
QC of daily composites	ocncolor*_int_dy	ocncolor*_dy	3.0	36.00	NA	108
	ocncolor*_int_3wk					
Form weekly QCed composite	ocncolor*_dy	ocncolor*_wk	3.3	5.14	NA	120
Form weekly QCed reference field	ocncolor*_wk	ocncolor*_3wk	103.3	5.14	NA	3720
MODIS Level 2 Production of Sea Surface Temperature and QA	MOD02_L1B	sst and sstqc	5.0	288.00	43726	1440
	MOD35_L2					
	MOD38_L2					
	MOD03_L1A					
	MOD28_L3_ref3W K					
	MOD04					

process ID	Input File IDs ocncolor*=lw,dr1 ,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	Output File IDs ocncolor*=lw,d r1,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	maximum minutes of CPU processin g time per execution	No. of Exec. /day	MB of output created by processes active every day	maximum minutes of CPU to process a day of data
	NMC or DAO					
MODIS SST Space Binning	sst	sst_**_int_p	0.5	288.00	4070.4	144.0
MODIS SST Production of Single Orbit	sst_**_int_p	sst_**_int_orb	0.1	48.00	2088.0	6.4
MODIS SST Daily Compositing of Data	sst_**_int_orb	sst_**_int_dy	4.0	4.00	2088.0	16.0
Form weekly composite (daytime data)	sst_**_int_dy	sst_**_int_wk	3.3	0.57	NA	13.3
Form weekly reference field (daytime data)	sst_**_int_wk	sst_**_int_3wk	103.3	0.57	NA	413.3
MODIS SST Level 3 QA for Daytime Data	sst_**_int_dy	sst_**_dy	3.0	4.00	NA	12.0
	sst_**_int_3wk					
Form weekly composite of QCed daily files	sst_**_dy	sst_**_wk	3.3	0.57	NA	13.3
Form weekly refernce of QCed data	sst_**_wk	sst_**_3wk	103.3	0.57	NA	413.3
MODIS Level 3 Daily Production of Oceans Productivity	MODOCCLR_L2	MOD27_L3_DY		355.00		
MODIS Level 3 Weekly Production of Oceans Productivity	MOD27_L3_DY	MOD27_L3_WK		50.71		
MODIS Level 3 Yearly Production of	MOD27_L3_WK	MOD27_L3_YR		0.97		

process ID	Input File IDs ocncolor*=lw,dr1 ,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	Output File IDs ocncolor*=lw,d r1,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	maximum minutes of CPU processin g time per execution	No. of Exec. /day	MB of output created by processes active every day	maximum minutes of CPU to process a day of data
Global Oceans Productivity						
SST MODIS Level 2 subsetting according to buoy locations. Output shipped to SCF.	MOD01_L1A	MOD28_L2_BUOY		576.00	21556.8	
	MOD02_L1B					
	MOD03_L1A					
	MOD35_L2					
	MOD38_L2					
	MOD28_L2					
	MOD04					
	ANC_AERTAU_THIR					
	ANC_BUOY_LOCATION S					
	ANC_ESI					
	DAS_SLU					
	DAS_SLV					
	DAS_WINDS					
	DAS_PSFC					
	DAS_SLP					
	DAS_PRSE					
	DAS_PRS_SPHU					
	DAS_PRS_TMPU					
	ANC_NMC_SURF					
	ANC_OP					
	ANC_SMSST					
	ANC_SMWLR					
Ocncolor MODIS Level 2 subsetting according to buoy locations. Output shipped to SCF.	MOD01_L1b	MOD18_L2_BUOY		288.00	4464.0	
MODIS Level 2 Subsetting of global 4-	MOD01_L1A	GLOBSUBS_L2		585.00		

process ID	Input File IDs ocncolor*=lw,dr1 ,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	Output File IDs ocncolor*=lw,d r1,dr2, qc ocncolor*_ = name for each of the 36 ocean color products sst_**=d or n	maximum minutes of CPU processin g time per execution	No. of Exec. /day	MB of output created by processes active every day	maximum minutes of CPU to process a day of data
km data.						
	MOD02_L1B					
	MOD03_L1A					
	ANC_AERTAU_THIR					
	MOD28_L2					
	MOD0CCLR_L2					
	MOD38_L2					
	ANC_ESI					
	DAS_SLU					
	DAS_SLV					
	DAS_WINDS					
	DAS_PSFC					
	DAS_SLP					
	DAS_PRSE					
	DAS_PRS_SPHU					
	DAS_PRS_TMPU					
	ANC_NMC_SURF					
	MOD35_L2					

3.2.2.4 Overall ocean processing time for the Level-2 radiances

The processing estimates presented in to previous section utilized a test data set that assumes no clouds or land. In addition, the at-sensor radiances/reflectances presently are not well mated to the atmospheric correction algorithms and consequently do not take advantage of provisions made in the codes to minimize processing time by utilizing pixel to pixel coherence in choice of potential atmospheric correction models. The Level-2 ocean color processing times could thus be overstated by as much as a factor of three. The Ocean Team is working the MODIS team to produce an improved test data set.

Our goal is to improve processing capacity through a combination of algorithm efficiency improvements and upgraded hardware to achieve a capability to process reduced resolution (4km) MODIS observation at a rate that exceeds 10 data days/day. This capability will permit changes such as algorithm enhancements or updated

calibration to be easily verified using long time series of global observations (> several years) prior to submitting the updates to the MODIS team for consideration. In particular, if we can process high resolution global data at a rate of 1 data day/day, we will meet the goal to process 4km reduced resolution data at a rate > 10 data days/day. Finally the algorithms are being reviewed to determine where computation efficiencies can be achieved.

3.2.2.5 Data Volume

Table 12 presents a list of the products that will be produced by the Level 2 and Level 3 PGEs for ocean color and SST. Product names, file status, and area of application and file size are given in the table. Interim files are expected to have a lifetime of less than a day.

Table 12. File sizes for SST and ocean color products.

File ID	Instrument	File Name	SPSO Equivalent	File Disposition	File Size (MB)
ocncolorlw	MODIS (AM)	MODIS Level 2 Ocean color algorithms	MOD18 MOD37 MOD39	archive	259.97
ocncolordr1	MODIS (AM)	MODIS Level 2 Ocean color algorithms	MOD19 MOD20 MOD23 MOD23 MOD25 MOD26 MOD31	archive	238.30
ocncolordr2	MODIS (AM)	MODIS Level 2 Ocean color algorithms	MOD21 MOD22 MOD24 MOD36	archive	281.63
ocncolorqc	MODIS (AM)	MODIS Level 2 Ocean color algorithms	none avail	archive	195.08
ocncolor*_int_p	MODIS (AM)	MODIS Ocean Color Space Binned Data	none avail	Interim	6.50
ocncolor*_int_orb	MODIS (AM)	MODIS Ocean Color Single Orbit	none avail	Interim	43.50
ocncolor*_int_dy	MODIS (AM)	MODIS Ocean Color Daily Composite	none avail	Interim	522.00

File ID	Instrument	File Name	SPSO Equivalent	File Disposition	File Size (MB)
ocncolor*_int_wk	MODIS (AM)	MODIS Ocean Color Weekly Composite	none avail	Interim	522.00
ocncolorl*_dy	MODIS (AM)	MODIS Ocean Color QCed Daily Composite	MOD18 -26 MOD31 MOD36 MOD37 MOD39	Archive	522.00
ocncolor*_wk	MODIS (AM)	MODIS Ocean Color QCed Weekly Composite	MOD18 -26 MOD31 MOD36 MOD37 MOD39	Archive	522.00
ocncolor*_int_3wk	MODIS (AM)	MODIS Ocean Color Weekly Reference Field	none avail	Interim	522.00
ocncolor*_3wk	MODIS (AM)	MODIS Ocean Color Weekly QCed Reference Field	MOD18 -26 MOD31 MOD36 MOD37 MOD39	Archive	522.00
MOD18_L2_BUOY	MODIS (AM)	MODIS Ocean color Product Data subetted for QA Work	Subset of MOD18	Interim	19.58
MOD28_L2_BUOY	MODIS (AM)	MODIS Sea Surface Temperature Product Data subetted for QA Work	Subset of MOD28 (2527, 5359)	Interim	9.08
sst	MODIS (AM)	MODIS Level 2 Sea Surface Temperature	MOD28 (2527, 5359)	archive	43.33
sstqc	MODIS (AM)	MODIS Level 2 Sea Surface Temperature QA	MOD28 (2527, 5359)	archive	108.50
sst_d_int_p sst_n_int_p	MODIS (AM)	MODIS Sea Surface Temperature Space Binned Data	MOD28 (2527, 5359)	Interim	13.01
sst_d_int_orb sst_n_int_orb	MODIS (AM)	MODIS Sea Surface Temperature Single Orbit (day & night)	MOD28 (2527, 5359)	Interim	75.40
sst_d_int_dy sst_n_int_dy	MODIS (AM)	MODIS Sea Surface Temperature daily Composite (day & night)	MOD28 (2527, 5359)	Interim	1044.00
sst_d_int_wk sst_n_int_wk	MODIS (AM)	MODIS Sea Surface Temperature Weekly Composite (day & night)	MOD28 (2527, 5359)	Interim	1044.00

File ID	Instrument	File Name	SPSO Equivalent	File Disposition	File Size (MB)
sst_d_dy sst_n_dy	MODIS (AM)	MODIS Sea Surface Temperature Daily Product QCed	MOD28 (2527, 5359)	Archive	1044.00
sst_d_wk sst_n_wk	MODIS (AM)	MODIS Sea Surface Temperature Weekly Product QCed	MOD28 (2527, 5359)	Archive	1044.00
sst_d_int_3wk sst_n_int_3wk	MODIS (AM)	MODIS Sea Surface Temperature Weekly Reference Field	MOD28 (2527, 5359)	Interim	1044.00
sst_d_3wk sst_n_3wk	MODIS (AM)	MODIS Sea Surface Temperature Weekly Reference FieldQCed	MOD28 (2527, 5359)	Archive	1044.00
MOD27_L3_DY	MODIS (AM)	MODIS Level 3 Daily Oceans Productivity	MOD27	Archive	3.97
MOD27_L3_WK	MODIS (AM)	MODIS Level 3 Weekly Oceans Productivity	MOD27	Archive	0.10
MOD27_L3_YR	MODIS (AM)	MODIS Level 3 Yearly Global Oceans Productivity	MOD27	Archive	0.10
Anc_AerTau_ThIR	Other/TBD	Data Set needed for Atmospheric Correction of MODIS ThIR Radiances	N/A	Interim	5.00
ANC_MDB_LOCATIONS	Other	Ancillary_Buoy Data	N/A	Permanent	2.00
GLOBSUBS_L2	MODIS (AM)	MODIS Ocean Color and SST Matchup Database of Subsetted MODIS Data (Global 4km Database)	N/A	Archive	21.88

3.2.2.5.1 Volume of data to be transferred to Miami

Data volumes that will be required in Miami have been computed using various approaches. The first involves requesting data for the matchup data base and is based on number of buoys, area to be subset, frequency of observation and includes the corresponding ancillary data. This estimate yields approximately 3 gigabytes/day for ocean color and 6 gigabytes/day for SST for near-real time buoy data availability (approximately 20% of the potential buoy observations). Increasing the buoy count to include times and locations of eventual buoy reports will increase the data volume by a factor of 5. Another approach assumes that the DAAC environment will not have the necessary support tools at launch which will require a subsetting program to be added to the ocean PGE with a resulting resolution of 4-km. At this resolution approximately 10 gigabytes/day will be required.

3.2.2.5.2 Total volume

Total volume requirements will range between 10 and 50 gigabytes/day depending on available DAAC system services.

3.2.2.6 Network Capacity

The previous analysis suggests 50 gigabytes will be exchanged each day which would require a sustained data rate of approximately 5 megabits/sec. Assuming a link efficiency of 25% to 50%, a link speed of 10 to 20 megabits/sec is necessary. The link to Miami will provide data for three MODIS investigators: Gordon, Brown and Evans. The link will also support access to the TLFC and for exchange of data and products between the DAAC, TLFC, MOTCF and other ocean SCF's. A DS3 (45 megabits/sec) or OC3 (155 megabits/sec) circuit is requested to support this network bandwidth.

3.2.3 Calibration Validation

3.2.3.1 Introduction

The matchup database provides a basis for comparison of coincident satellite and *in situ* observations. Experience gained from application of this type of data during the CZCS mission is presented as a guide for a MODIS strategy. The procedures presented below will be modified to reflect experience gained from SeaWiFS where both the *in situ* (buoy and ship) and satellite sensor data closely reflect the data flow that will be available for MODIS. Data included in the *in situ* portion of the matchup database will be calibrated and validated by the data suppliers. The calibration time series will be compared with similar calculations produced by the MCST using the MODIS on board calibrators. Finally the time series will be referenced to the lunar and solar time series produced by the MCST.

The vicarious method will be augmented by using space and time series comparisons assembled using the 4km reduced resolution global and selected area 1km data sets to enable testing of long term calibration trends and evaluation of new algorithms to be performed. As CPU performance increases, we expect to be able to run data at rates of 20-100 data days/day on global data sets and thus have the ability to easily produce extensive time series that will form the basis for comparison of effects of application of calibration or algorithm changes. These data sets would be available to test other algorithms dependent on MODIS ocean water leaving radiances.

3.2.3.2 Satellite Field Analysis Techniques

Comparisons to assess algorithm performance are not limited to results obtained using the MDB. The following sequence of figures demonstrates a series of comparisons based on analyzing fields obtained from applying the geophysical algorithm to global fields. This sequence is based on processing AVHRR infrared retrievals for the period 1987 through 1993 as part of the AVHRR Ocean Pathfinder project. Use of global fields permits us to address some of the comments raised by the ATBD review where the reviewers recognized the limited temporal and geographical coverage afforded by the *in situ* data sets and their inability to completely address algorithm performance throughout the satellite data set.

A multi-year time series of satellite fields has been assembled and compared to itself in time as well as to a 'blended analysis' produced by R. Reynolds of the NOAA Climate Analysis Center. Reynolds uses a combination of ship and buoy *in situ* observations to

correct NOAA satellite observations using an Optimal Interpolation technique. The Pathfinder fields use only buoy observations to determine algorithm coefficients and no further constraints are applied. Due to the filters applied while constructing the MDB, approximately 90% of the buoy observations are removed from the MDB leaving the remaining buoy observations as an independent comparison data set. These observations are used to establish local, regional, and temporal algorithm behavior by using an Objective Analysis technique to compute complete fields (cloud fill) using only satellite data. The OA'ed fields are then compared with the buoy observations (not shown).

Figure 8 shows one week (1991-week12) of NOAA-11 Pathfinder SST mapped to a standard equal angle 1 degree grid. Figure 5 shows the corresponding field produced by the Reynolds blended analysis. These weekly fields form the basis for much of the analysis. Figure 8 shows coverage typical of a weekly period. Work with shorter time period fields must be able to deal with a nominal 25% useful areal coverage or requires some form of gap filling procedure such as the Reynolds Optimal Interpolation or Mariano's Objective Analysis programs. Figure 8 present the basic, non-interpolated, satellite retrievals for 1991, week 12. The field is composed of all ascending 4km AVHRR GAC (Global Area Coverage) that passed the cloud tests during the week. The data is assembled as a series of daily files and mapped to an equal area 9km bin projection. The daily files are summed in space and time to create the 1 degree equal angle weekly map corresponding to the Reynolds analysis.

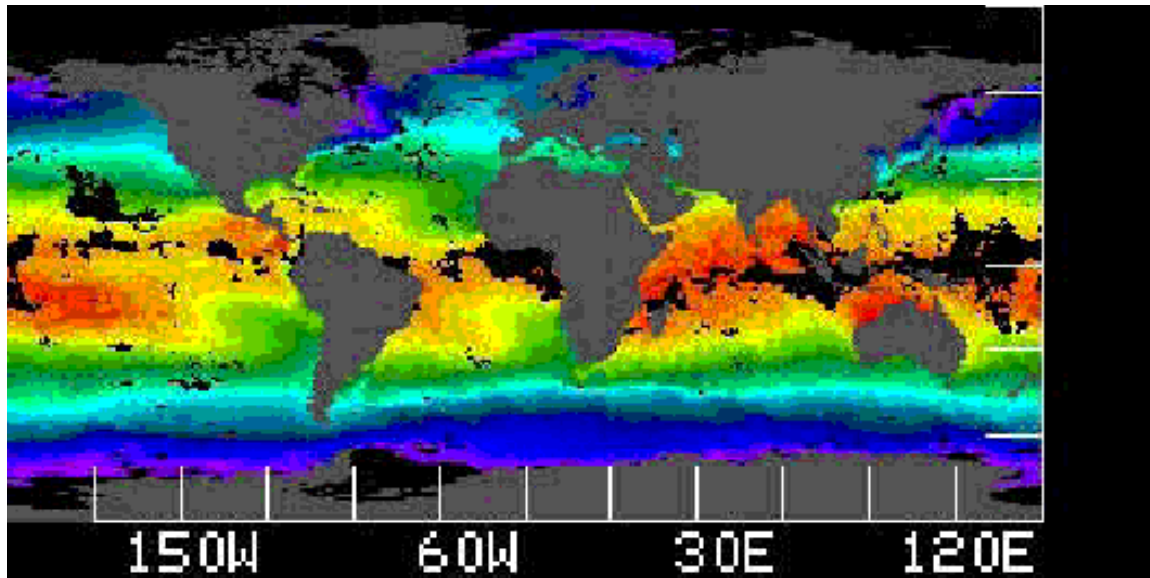


Figure 8. Pathfinder SST image for 1991, week 12. Black areas were continuously cloudy for the week; gray regions are land. Blues areas are cold SST while red regions are warm.

The Reynolds blended analysis, Figure 9, is generated from NOAA satellite retrievals for the week. The retrievals are sorted into 1 degree bins and averaged. An Optimal Interpolation step that incorporates ship and drifting buoy data is used to correct the satellite retrievals and interpolate through regions that had no retrievals. Polar region and area south of the equator are sparsely represented in the ship data and are more subject to biases introduced by a particular ship or represent only the NOAA satellite data.

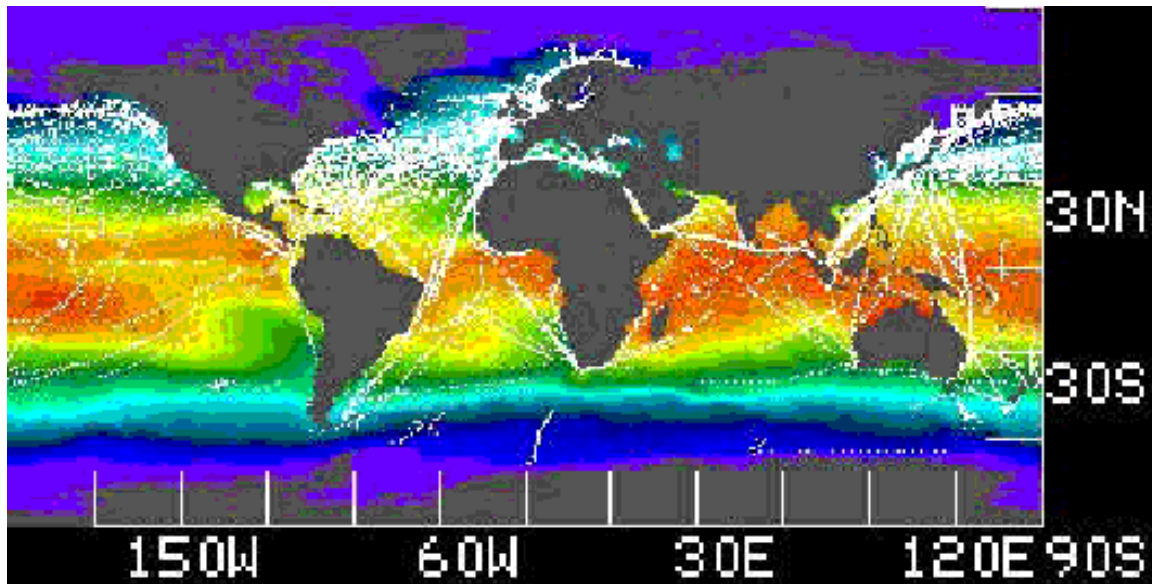


Figure 9. Reynolds blended analysis for 1991, week 12. White areas are locations of ship observations used as the reference for the Optimal Interpolation correction of the NOAA SST retrievals.

Figure 10 presents the difference map between the Pathfinder and Reynolds analysis for week 12 of 1991. The gray areas, differences $< 0.33^{\circ}\text{C}$, represent the regions of agreement between the two products. The regions where differences are larger are seasonal. Research is underway to try and establish how the differences are partitioned due to factors such as the process of developing each product, the nature of the satellite SST retrieval process and geophysical considerations such as ocean stratification, distribution of atmospheric water vapor, wind...

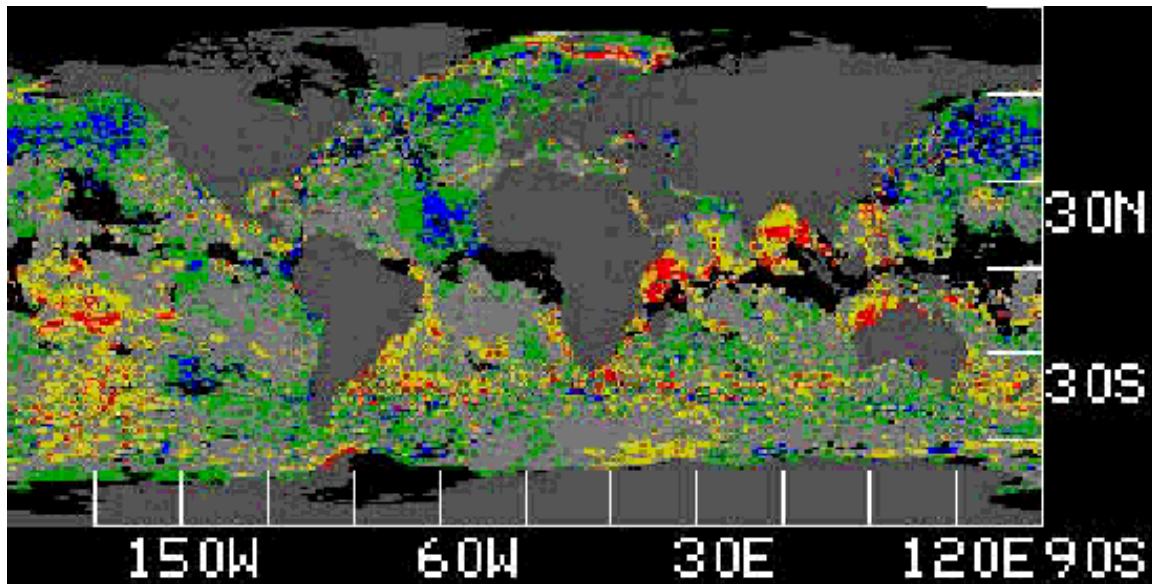


Figure 10. Pathfinder - Reynolds difference map. Light gray regions show areas where the Pathfinder and Reynolds fields differ by less than 0.33°C. Green and yellow areas are 0.33 to 1°C cold and warm respectively while the blue and red areas show cold and warm departures > 1°C. A significant fraction of the regions of large discrepancy lies in areas with little to no ship data.

3.2.3.3 Temporal comparison of temperature fields

Time provides another dimension to analyze the performance of the satellite retrieval algorithms. The interannual behavior of the Pathfinder-Reynolds differences seen in the following two plots (Figures 11 and 12) of one week for each of four years, 1990-1993. Figure 11 shows the zonally averaged differences for week 12. The small jumps seen in each of the curves in Figure 11 is 0.15°C or slightly more than the least count digitization for the long wavelength channels (10 and 11 μ) of the AVHRR sensor. For most of the plot, data for the four years falls within ± 1 count quantifying the stability of the retrievals between the two products, each referenced to a different *in situ* data set. The systematic behavior for the four years results suggests other sources of SST (e.g. wider geographical distribution) will be required to address these differences.

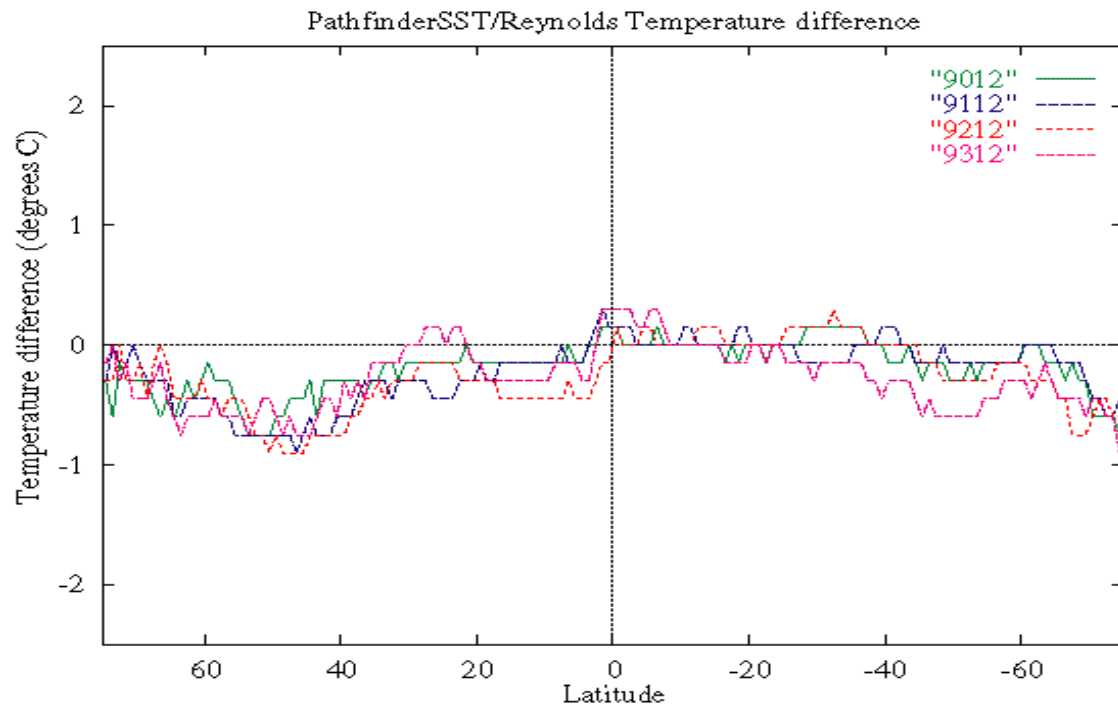


Figure 11. Four year presentation of Pathfinder-Reynolds difference for zonally averaged week 12, years 1990 to 1993, NOAA-11. Small vertical excursions in the plot are at the digitization limit of the AVHRR sensor. Temperature patterns for the four years do not show temporal drift. Consistent differences between the Pathfinder and Reynolds fields will be analyzed and compared with other data sets such as wind, water vapor, sensor state, *in situ* observations to help determine the source of the discrepancies.

Figure 12 shows the zonally averaged differences for week 38. The curve for 1991 shows the latitudinal range influenced by the Mt. Pinatubo aerosols, 30N to 25S. Again data for the four years is tightly grouped with more agreement seen in the northern hemisphere than in the southern. The peak of the aerosol influence is seen at 10N, the latitude of the volcano.

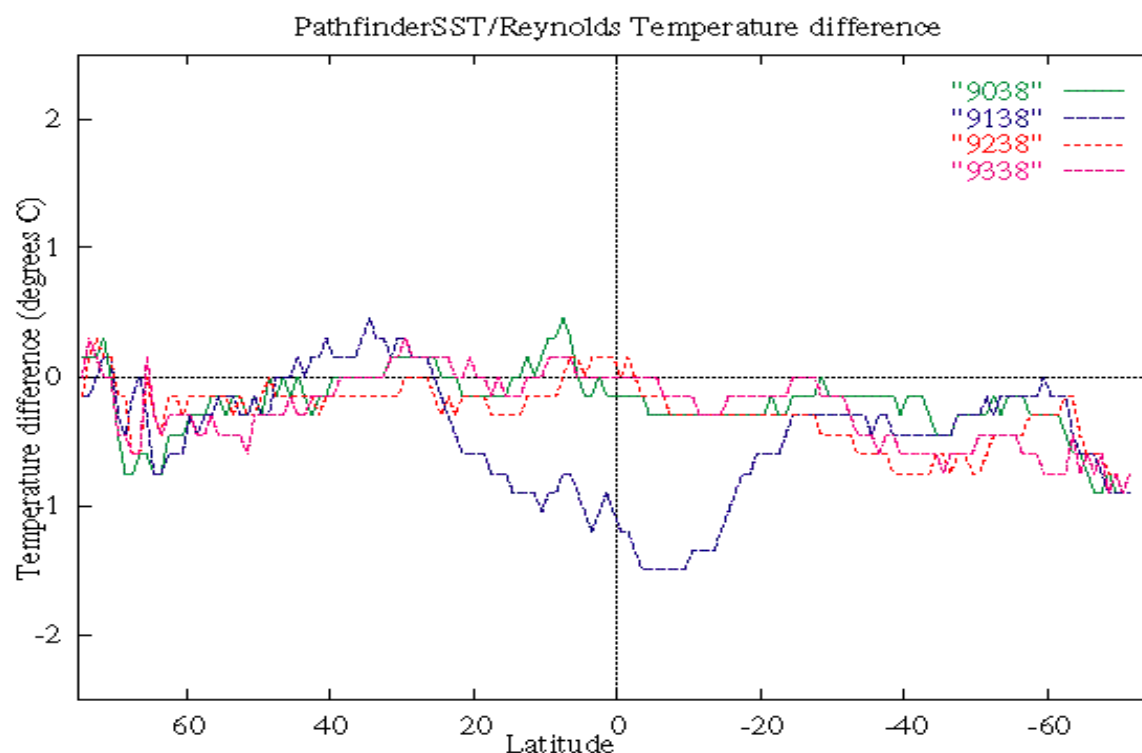


Figure 12. Same presentation as Figure 11 for week 38 showing the effect of the Mt. Pinatubo aerosols on the Pathfinder algorithm in 1993

3.2.4 Quality Control and Diagnostics

3.2.4.1 Introduction

A number of automated quality control mechanisms have been included in the Pathfinder SST processing path. During Level-2 calculation a series of tests is employed that utilize spatial and spectral properties to determine suitability of the sensor retrieval. The calculated SST is next compared with a climatology local to the time period, *e.g.* a SST field produced from an OA analysis of the last week's SST. Finally a longer space-time series is used to produce a better estimate of the expected SST and its variance. Results from each of the tests are used to set a flag. A rule base has been formulated for AVHRR Pathfinder that uses the flags, pointing angles, and SST variance to determine the expected quality of a given retrieval. These tests have been employed in the Pathfinder processing and have resulted in approximately 25% of the satellite observations classified as good and the remainder placed in a lower quality category. Of the 2×10^7 retrievals calculated for each SST field, approximately 100 bad retrievals

are classified as good. Further study is required to determine the size of the good fraction that is misclassified as bad. Each pixel in the result's field has a companion status field. The status field is used in subsequent processing such as preparing daily to longer term global maps. Each map location contains data ranging in quality from bad to good permitting location of persistent cloud contaminated areas.

Additional tests can be performed such as generating difference maps of daily field - reference climatology, zonal and meridional sections. Each map location will have an expectation of acceptable variance. The processing software can perform these tests and report results to a Q/A monitor. The expected outcome of a discrepancy is a report to the SCF for further analysis. These tests provide quality testing for mapped global fields.

The time series generated from the matchup data base also provides a quantitative test on the combined sensor and algorithm performance. These tests provide only a limited analysis of product and sensor stability when utilized on a single retrieval (day) basis. Results obtained from analysis of longer time series will be required to differentiate the various failure modes. For example, the sensor could undergo undiagnosed calibration shifts with time or the algorithm's correction for scan angle might be incomplete. A suitable number of samples is required to develop a suitable set of tests given the expected retrieval noise.

Validation of MODIS retrievals will be based on comparisons with *in situ* data obtained from ship and buoy observations. Assuming continuation of the present international buoy deployment efforts, order of 500-1000 drifters should be available to provide reference for SST. Brown intends to augment this data set with a number of ship base radiometers. Clark will deploy optical buoys and we expect that several investigators will deploy optical drifters. These measurements provide absolute or relative reference data depending on the ability to establish and maintain calibration. In addition, comparison will be made using consistency between fields in space and time. Fields derived from other satellite sources (AVHRR, ATSR, OCTS, SeaWiFS, multiple MODIS instruments) provide additional comparison opportunities.

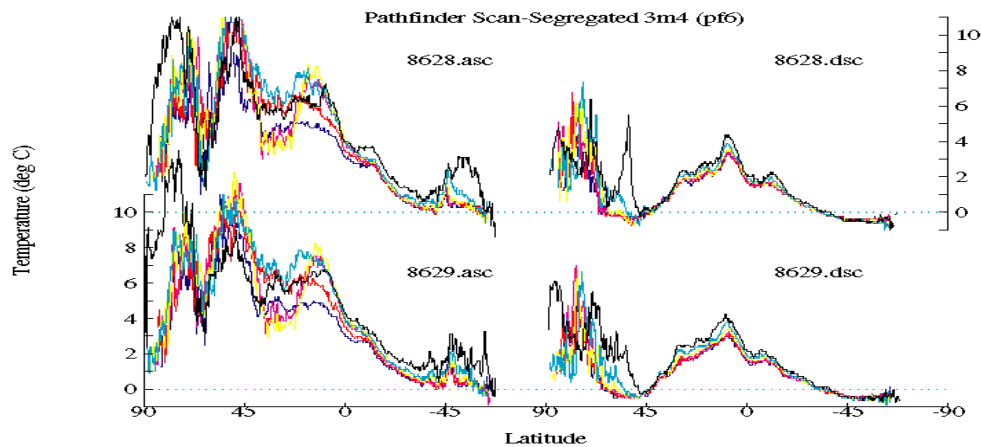
The sensor calibration time series obtained from the vicarious method will be compared with the record developed by the MCST. The calibration record developed by MCST using MODIS sensor sources will be the only method of tracking orbit to orbit sensor

stability. The vicarious and matchup database will provide long term trend analysis and calibration, validation for the retrieved water leaving radiances for time periods longer than ~six months. Products produced at the MOTCF for a subset of the MODIS observations will be compared with the standard products produced at the PGS to verify proper operation of the algorithm.

3.2.4.2 Potential quality effects of stray light

An important application of the quality control analysis is to understand the range of observation conditions that yield useful surface retrievals. Figures 13a and b are based on first producing global day and night fields of various satellite measurements. In the following Figures 13a and b, the channel 3 - channel 4 (MODIS channels 20 and 31) difference is calculated and the results stratified by 10 degree satellite zenith angle bins. Daily fields are produced and subsequently combined into weekly fields. The weekly fields are then zonally averaged to show variation in the channel 3-4 with latitude for selected weekly periods. Each of the figures presents daytime (ascending on left) and nighttime (descending on right) with satellite zenith angle range discriminated by color. Channel 3, 3.7 μ band, contains both emissive and reflective solar energy during daytime observations increasing the channel 3-4 difference. Nighttime retrievals should show only the effects of radiance emitted from the ocean surface and intervening atmosphere. The large spikes that occur in the northern latitudes (poleward of 45 degrees) during mid-year (week 28) for large satellite zenith angles (>50 deg) suggest that energy is being received by the detectors from sources other than the surface. Later in the year for week 41, smaller abrupt transitions are seen for angles >40 deg. In each case we would need to restrict retrievals for time and locations that are not subject to these problems. An added benefit is the production of a climatology for the channel difference that permits quality testing of the individual retrievals to determine whether a given retrieval is valid.

a.



b.

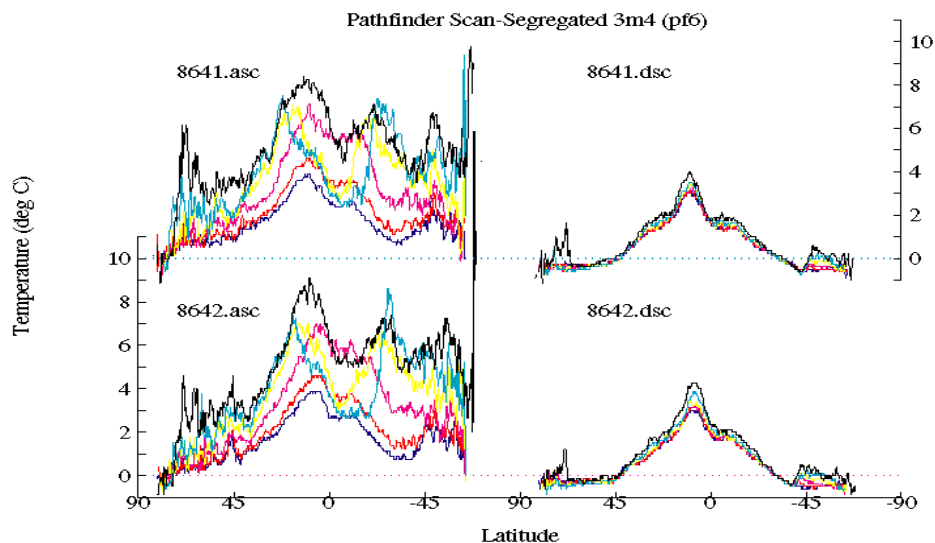


Figure 13. Pathfinder channel 3 - channel 4 differences (MODIS Ch20-Ch31) for two 1 week period, 1986 week 28, upper panels, and week 41, lower panels. Day orbits presented on left(asc) and night orbits on right (dsc). Data processed as daily, 9km, equal area fields, time binned into 7 day week periods, and then zonally averaged. Colors represent satellite zenith angle ranges, dark blue, 0-10 deg; red, 10-20; pink 20-30, yellow 30-40; light blue 40-50; black >50. Large differences for day passes shows influence of reflected sunlight in channel 3. Abrupt increases in channel difference poleward of 45 deg at high satellite zenith angles suggests stray light problems. Finally increasing channel difference near the equator reflects 'average' effects of path radiance and serves as a baseline for possible algorithm option selection and for data quality tests.

3.2.4.3 Quality control of *in situ* observations

A fundamental basis of both the SST and ocean color calibration and validation activity is based on comparison of satellite retrievals with corresponding *in situ* observations. Since the *in situ* observation are subject to error and calibration drift, especially for drifting buoys, we have developed a preliminary approach to identifying buoys with potentially faulty measurements. Figures 14 and 15 are based on processing satellite retrievals using the equations and quality filtering based on the previously presented matchup process. The resulting daily satellite fields are then Objectively Analyzed to remove data gaps due to cloud or instrument problems. Residuals are computed for the individual OA Satellite SST - buoy SST and histogrammed by year.

In Figure 14, histograms are presented showing a gaussian-like distribution of residuals principally centered between $\pm 1\text{C}$ with biases $< 0.2\text{C}$. Neglecting partitioning the residuals by various criteria, e.g. satellite and solar zenith angle, time, location, ancillary data..., the histograms present an indication of the performance of the algorithm.

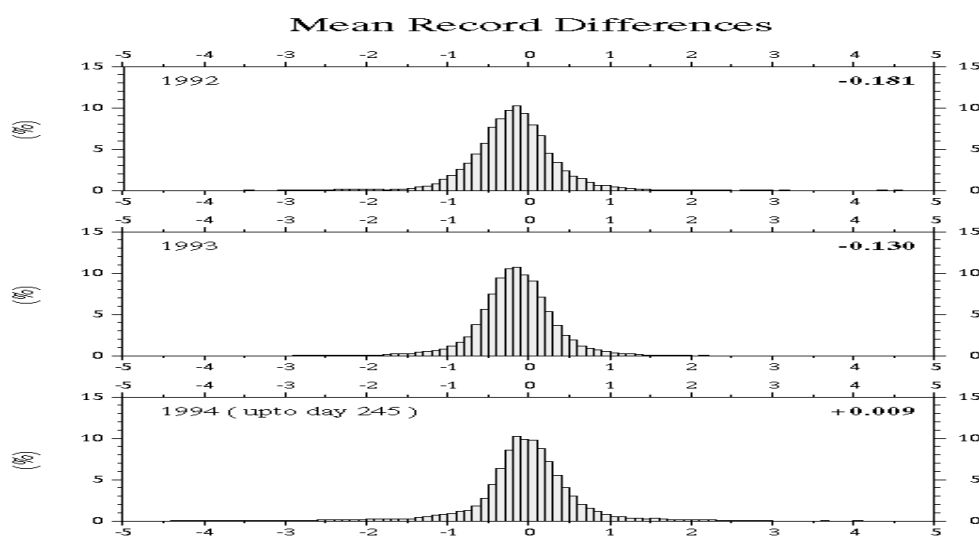


Figure 14. OA Pathfinder SST - Buoy SST for all buoys for 1992 to 1994. Daily Pathfinder SST fields are Objectively Analyzed to fill data voids (clouds).

Figure 15 restructures the data by computing residuals on a buoy by buoy basis. Here we assume that the satellite SST field provides a transfer reference between buoys in similar oceanographic and atmospheric conditions. The long term trends in the individual buoy residual also can be analyzed to determine whether a constant offset

(bias) is present, the residuals are distributed about zero or begin to depart at some point during the buoy lifetime. Most average residuals fall within $\pm 0.5^\circ\text{C}$ but some buoys report average temperatures that depart from the satellite field by more than 1°C . Buoys that exhibit this characteristic must be individually examined to determine if this behavior is due to buoy calibration or anomolous atmospheric conditions, e.g. unusual atmospheric moisture profiles or aerosols.

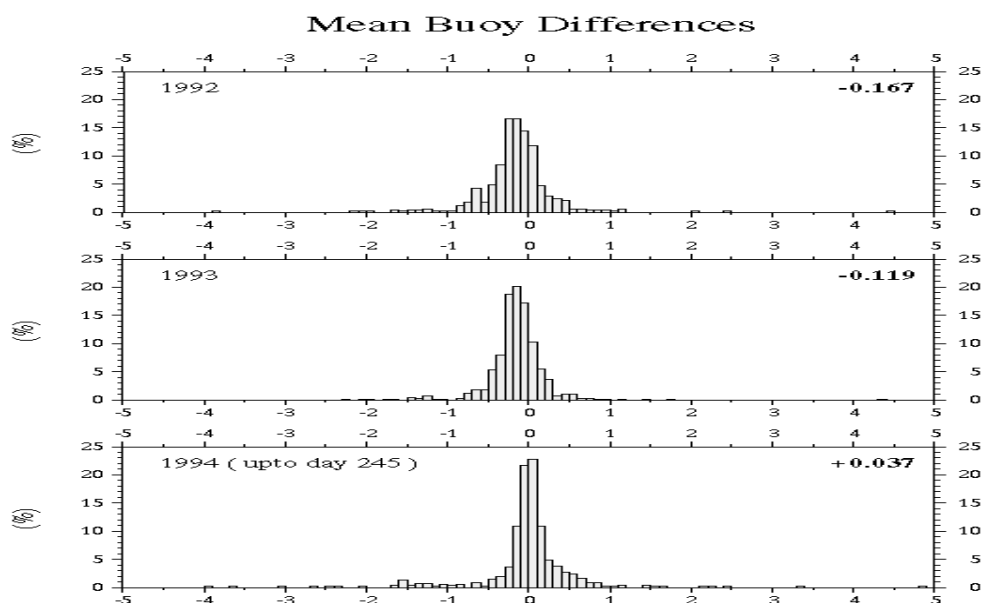


Figure 15. Mean OA Pathfinder SST - Buoy SST per buoy, histogrammed for all buoys for 1992 to 1994. Daily Pathfinder SST fields are Objectively Analyzed to fill data voids (clouds). Departures from 0 are indicative of possible buoy calibration error.

The next two figures (16 and 17) present a time series of the comparison of OA satellite SST and buoy SST. Figures 16 and 17 show the behavior of two buoys, one considered “good” and another “bad”. The four panels for each figure show from top down SST time series, satellite-buoy residuals by retrieval and with an 8 day filter to minimize effects of satellite zenith angle, a histogram of the residuals for this buoy and finally the trajectory of the buoy. This particular buoy (Fig. 16) compares well with the satellite SST for most of the record. The record also helps localize the time and location of maximum residual, near day 180 and thus points to the specific satellite fields that can

be further analyzed to determine if the buoy is located near a front or eddy and thus subject to errors introduced by navigation error.

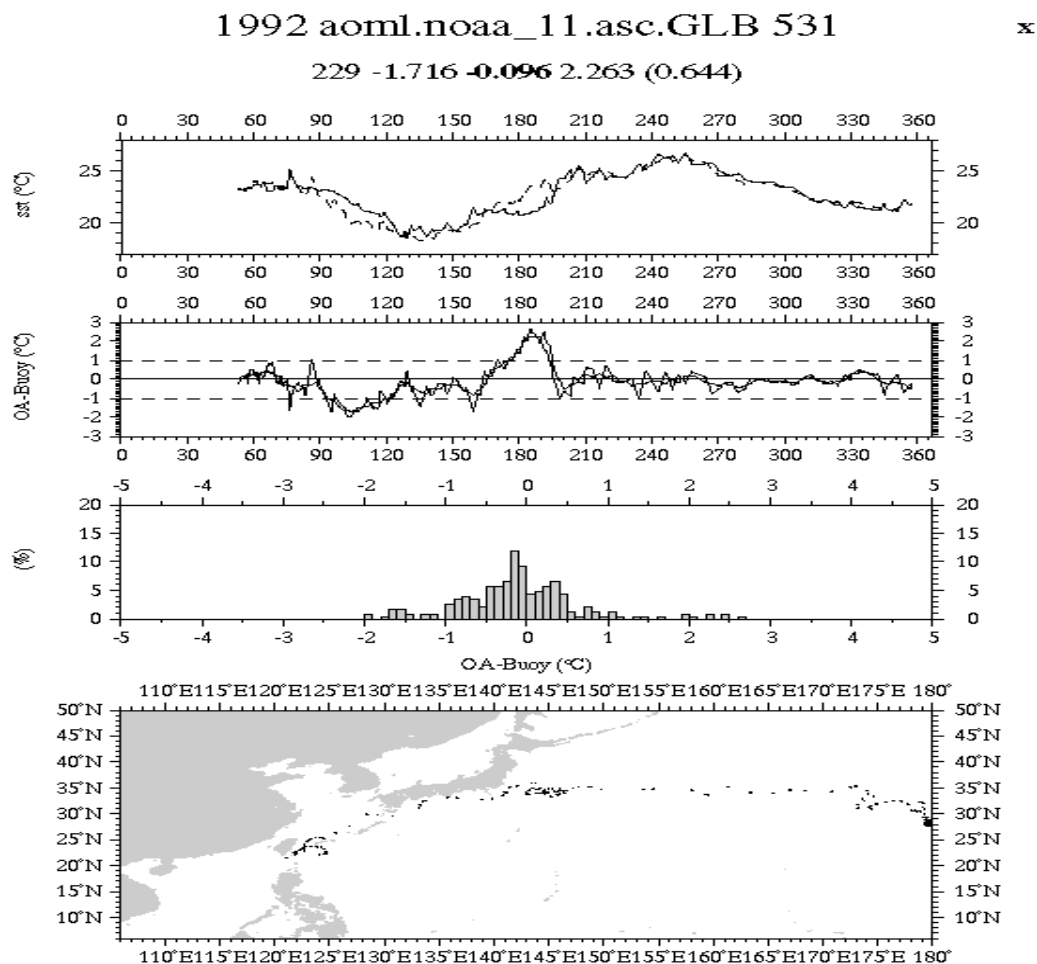


Figure 16. Plot of Buoy and OA Pathfinder SST, upper panel; OA Pathfinder - buoy SST, middle upper panel, OA Pathfinder - Buoy SST histogram, lower middle panel, and buoy trajectory, lower panel. Statistics for the lifetime of a 'good' buoy where residuals oscillate about the zero line for the life of the buoy.

Figure 17 shows a buoy where a consistent offset is seen throughout the buoy record. The type of trend is likely a result of buoy miscalibration. Care must be taken to insure that the offsets are not an indication of presence of aerosols or other atmospheric conditions. Correlation with other satellite or ancillary fields can assist in this process.

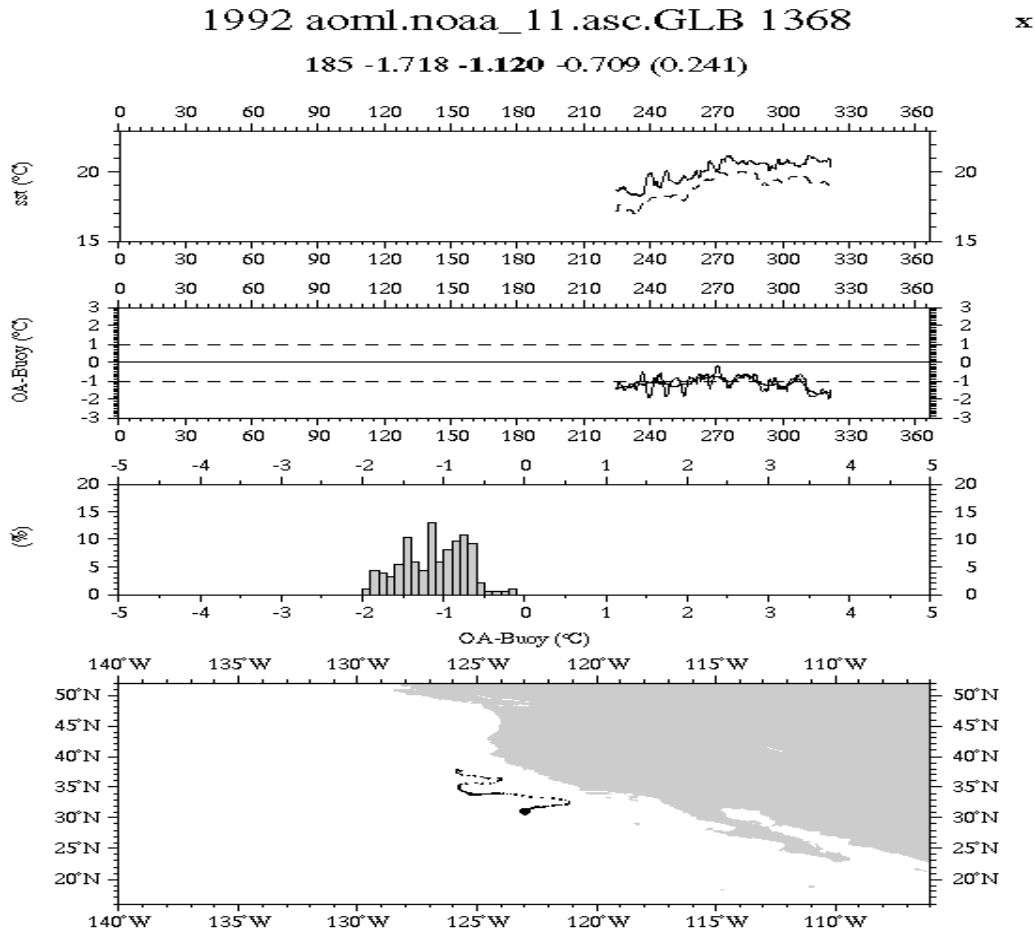


Figure 17. Plot of Buoy and OA Pathfinder SST, upper panel; OA Pathfinder - buoy SST, middle upper panel, OA Pathfinder - Buoy SST histogram, lower middle panel, and buoy trajectory, lower panel. Statistics for the lifetime of a 'bad' buoy where residuals oscillate about the zero line for the life of the buoy. Consistent offset of the OA Pathfinder SST and the buoy SST indicate a buoy calibration (bias) problem or possible unusual atmospheric condition.

3.2.4.4 Ocean Processing Level-2 Quality control flags

A series of quality flags have been defined to aid in the interpretation and selection of data. The flags indicate ranges of observational parameters where product calculation could be subject to increased error, e.g. large satellite or solar zenith angles, atmospheric conditions near the limit of retrieval capability, e.g. proximity to clouds, high aerosol or water vapor loading or specific in water conditions, e.g. low chlorophyll leading to low fluorescence, coccolithophorid blooms with high in water scattering. By

testing combinations of conditions, the suitability of a particular product retrieval for use in qualitative or quantitative applications can be determined.

The matchup database has been used for SST to define tests to help differentiate clear and cloudy pixels. When the additional spectral channels provided by MODIS become available, it is hoped that further tests can be defined that will help identify pixels subject to aerosol.

3.2.4.4.1 SST atmospheric correction processing flags

MODSST flag names:

define B_SatZTemp	1	# satellite zenith angle and temp test
define B_6_7_Diff	2	# 31/32 difference
define B_5_7_Diff	3	# 20/32 difference
define B_6_7_Max	4	# 31/32 max
define B_6_7_Unif1	5	# 31/32 uniformity test 1
define B_6_7_Unif2	6	# 31/32 uniformity test 2
define B_Hi_Sat_Zen	7	# sat zenith > LIMIT
define B_Ref_Thresh	8	# sst diff from reference
define B_1_Unif	9	# channel 20 uniformity
define B_2_Unif	10	# channel 22 uniformity
define B_5_Unif	11	# channel 23 uniformity
define B_Sst_Bad	12	# sst out of range
define B_5_6_7_Thresh	13	# 20,31,32 thresh
define B_Ascending	14	# ascending
define B_No_Data	15	# no data
define B_Glint	16	# glint

3.2.4.4.2 Additional SST Flags To Be Added

Recent instrument characterization results suggest MODIS instrument specific tests that will augment the algorithm specific tests. Both the need for such tests and the specific nature of the tests can be determined only after the MODIS instrument characterization is completed and the data studied by the MCST. Implementation of these tests will occur by late Spring or early Summer of 1997. These flags will be applied at the band and pixel level. Additional flag topics presently under consideration:

- Polarization within limits
- Spectral cross talk problem

Spatial cross talk problem

- potential near field scattering

- potential far field scattering

Lw < 0 or not correctable - Lt out of range

Additional bits for instrument/band specific problems that emerge from testing

Time histories and field analyses and climatologies will be employed once the data becomes available. Some of the fields initially can be populated using results from heritage instruments while others must await the launch of MODIS. Two examples are included:

Channel calibration changes - compare with history

- | Lw-reference | out of range

Two additional spectral test are being implemented for the SST product:

Channel 22-23 difference

Channel 22-23 vs 31-32 difference

Cloud tests from the Menzel cloud product will be incorporated, these include:

Cirrus cloud test

Low level cloud or fog

We expect that once the combination of spectral channels on MODIS is available, a number of factors that influence both ocean color and SST retrievals can be tested. Expected examples include:

White caps

Absorbing aerosols

Atmospheric correction model carry over

Cirrus clouds

Delta sst vs delta chl comparison, possible absorbing aerosol test

3.2.4.4.3 Ocean color atmospheric correction processing flags

MODCOL Flags

Current flag bit definitions:

define B_Atmos_Corr	1	# invalid algorithm results
define B_Land	2	# land pixel
define B_Supp_Data	3	# Invalid support data
define B_Sun_Glint	4	# High sun glint
define B_Hi_Lt	5	# Lt > Lt(knee)
define B_Hi_Sat_Zen	6	# sat zenith > LIMIT
define B_Shallow	7	# shallow water
define B_Bad_Lw	8	# Any Lw < 0
define B_Ghost	9	# Withing GHOST_LEN_xx of Lt > Lt(knee)
define B_Cloudy	10	# Cloud detected (albedo > threshold)
define B_Cocco	11	# Coccolithophorid bloom detected
define B_TurbidCase2	12	# Turbid Case 2 water
define B_Hi_Sol_Zen	13	# solar zenith > LIMIT
define B_Hi_La865	14	# La865 > fraction of Lt865
define B_Lo_Lw550	15	# Lw550 < ?
define B_Chlor_Al	16	# Chlorophyll algorithm failure
define B_STRONGABSORB	1	#
define B_CLOUDBIT1	2	#
define B_CLOUDBIT2	3	#
define B_SUNGLINT2	4	#
define B_SEAICE	5	#
define B_FLHOUT	6	# FLH out of range
define B_FLHBELOW	7	# L748 > L667
define B_FLHBASE	8	# L678 peak below baseline
define B_CFENONE1	9	#
define B_CFENONE2	10	#
define B_CHLORAFAIL1	11	#
define B_CHLORAFAIL2	12	#
define B_EPSCLEARFAIL	13	#
define B_unused	14	#
#define B_unused	15	#
#define B_unused	16	#

Specific product flags defined by F. Hoge for PEB, PUB:

Input radiances valid

Specific product flags define by H. Gordon coccolithophore:

LoRadiance(integer) - value below lower bound of lookup table

HiRadiance(integer) - value above upper bound of lookup table
InvalidEntry(integer) - invalid data in lookup table

Specific product flags defined by K. Carder

CONDITION/FLAG	DESCRIPTION
CASE 2 CHLOROPHYLL	
Rrs at 412, 443, 488, 551 < 0	If any input Rrs value is neg., set flag, stop processing.
Rrs412 < 0.0008	If Rrs412 is lower than threshold value, set flag, continue processing.
Rrs551 < 0.0010	If Rrs555 is lower than threshold value, set flag, continue processing.
chl_mod/chl_def < factor of 2	Check if semi-analytical chl is within factor of 2 to default chl, set flag.
default chl	If semi-analytical algorithm fails to return a value, use default chl algorithm, set flag.
chl quality	Chl quality level. TBD.
high scattering	High scattering flag. TBD.
CLEAR-WATER EPSILON	
chl_def < 0.5 mg/m3	Check if pixel is in a "clear-water" region.
e(531,667) < 0.95	If e(531,667) less than threshold, set flag. Possible iron-rich aerosols or dust.
IPAR and ARP	
default chl	Check if chl algorithm used default chl value.

Specific product flags defined by M. Abbott for CFE

CFE output flags description (First 2 bit flag) CFE_1

- | | |
|--|----|
| 1) FLH input warnings | |
| FLH_1 = 11 or 10 | 10 |
| FLH_1 = 01 | 01 |
| 2) FLH below expected range | 10 |
| 3) FLH above expected range | 10 |
| 4) Wrong baseline slope | 01 |
| 5) High coefficient of variation in 5 by 5 box | 01 |

Number of pixels used to calculate FLH (Second 2 bit flag) CFE_2

- | | |
|----|------------------------|
| 00 | 1 pixel used |
| 01 | 2-8 pixels used |
| 10 | 9-15 pixels used |
| 11 | 16 or more pixels used |

CFE output flags description

- | | |
|--|--|
| 1) ARP chlorophyll algorithm | CFE_3 = 0 (algorithm 1 was used)
CFE_3 = 1 (algorithm 2 was used) |
| 2) ARP "seasonal" range for absorption | CFE_4 = 0 (seasonal range 1 was used)
CFE_4 = 1 (seasonal range 2 was used) |
| 3) ARP low Lw(443) | CFE_5 |
| 4) ARP low Lw(412) | CFE_5 |
| 5) ARP high suspended sediments | CFE_5 |
| 6) ARP coccolith blooms | CFE_5 |
| 7) ARP shallow water | CFE_6 |
| 8) ARP algorithm failure | CFE_6 |
| 9) CFE below expected value | CFE_7 |
| 10) CFE above expected value | CFE_8 |

CFE binning

Pick average of "best" pixels within time/space bin.

Do not include in bin if: CFE_1 = 10 or 11
or CFE_6 = 1
or CFE_7 = 1
or CFE_8 = 1

Order of "best" pixels to include in bin. Only include pixels of equal ranking in binning. That is, if one pixel in the bin is in Rank #1,

include only this pixel in the bin. If there are no pixels in Rank #1, then include only pixels of Rank #2 and so on.

Specific product flags define by M. Abbott for Fluorescence Line Height (MOD23)

Input Data

Water-leaving radiance MOD18

Bio-optical algorithms MOD19

Expected input flags

BD_1 atmos. algor. failure

BD_2 land

BD_3 invalid support data

BD_4 sun glint

BD_5 (not used)

BD_6 sat. zenith

BD_7 shallow water

BD_8 Lw < 0

BD_9 ghosting

BD_10 cloud

BD_11 coccoliths

BD_12 Case 2 water

BD_13 solar zenith

BD_14 La(865) high

BD_15 Lw(550) < min.

BD_16 chlor. algorithm failure

FLH output flags description (First 2 bit flag) FLH_1

1) BD_1 Atmos. algor. Failure	11
2) BD_2 Land	11
3) BD_3 Invalid support data	01
4) BD_4 Sun glint	11
5) BD_5 (not used)	
6) BD_6 Sat. zenith	10
7) BD_7 Shallow water	11
8) BD_8 Lw < 0	11
9) BD_9 Ghosting	10
10) BD_10 Cloud	11
11) BD_11 Coccoliths	01
12) BD_12 Case 2 water	01
13) BD_13 Solar zenith	11
14) BD_14 La(865) high	11
15) BD_15 Lw(550) < min.	10

16) BD_16 Chlor. algor. Failure 10

FLH output flags description (1 bit flags) Flag number

- | | |
|-----------------------------|-------|
| 1) FLH below expected range | FLH_2 |
| 2) FLH above expected range | FLH_3 |
| 3) Wrong baseline slope | FLH_4 |
| 4) FLH below baseline | FLH_5 |

Number of pixels used to calculate FLH (2 bit flag) FLH_6

- 00 1 pixel used
- 01 2-8 pixels used
- 10 9-15 pixels used
- 11 16 or more pixels used

FLH output flag description (1 bit flag)

- | | |
|--|-------|
| 1) High coefficient of variation in 5 by 5 box | FLH_7 |
|--|-------|

FLH binning

Pick average of "best" pixels within time/space bin.

- | | |
|---------------------------|------------------|
| Do not include in bin if: | FLH_1 = 11 or 10 |
| or | FLH_2 = 1 |
| or | FLH_3 = 1 |

Specific product flags defined by W. ESAIAS for productivity:

<u>CHLOROPHYLL FLAGS:</u>	# bits	position
Chlorophyll a Field Id (0-5)	4	0-3
Chlorophyll a Quality Index (0=Good,15=Bad,1-14=Ugly)	4	4-7
High-Variance Class	1	8
Low-Sampling Rate (n < _____)	1	9
Open Ocean Flag (1=ocean, 0=inland waters)	1	10
Coastline in Bin	1	11
Depth < 30 meters	1	12
High Average Aerosol	1	13
Low Average Radiance	1	14
High Cloud Frequency	1	15

(added to running annual average field in V2)

SUBPRODUCT FLAGS:	# bits	position
Function Id (0-15)	4	0-3
Quality Index (0=Good,15=Bad,1-14=Ugly)	4	4-7
No Chlorophyll a	1	8
No Chlorophyll a Uncertainty	1	9
No Sea Surface Temperature	1	10
No Sea Surface Temperature Uncertainty	1	11
No Mixed-Layer Depth	1	12
No Mixed-Layer Depth Uncertainty	1	13
No Average Daily PAR	1	14
No Average Daily PAR Uncertainty	1	15

3.2.5 Exception Handling

Exception handling for ocean processing is relatively simple, products are computed at each pixel if inputs are valid. Each investigator has defined specific tests appropriate to the algorithm. Results are produced and flagged as appropriate. The flags in turn are used when Level-3 products are produced. When multiple quality levels are present, pixels are binned at the highest available quality level and data at any lower quality levels are discarded for each pixel. Summary statistics are retained in the output granule to complement the pixel levels flags carried for each separate product.

It is expected that all ancillary data will be pre-screened prior to Level-2 calculation and alternatives selected if missing or erroneous data is detected. If substantial delay is experienced in receiving validated ancillary data, suitable climatologies will be used for initial reprocessing and a subsequent processing run will be required once proper ancillary data becomes available. SeaWiFS has scheduled four levels of reprocessing separate from calibration run to accommodate updates in ancillary fields, calibration and algorithm improvements. The time frame for these reprocessing efforts ranges from several days to a year or more.

3.2.6 Data Dependencies

Data dependencies have been minimized for pre-launch and for initial post-launch processing. All ancillary data used in processing is based on presently available NOAA or NASA data products. Transition to EOS products will be undertaken once the respective products have been validated and global products become routinely available in a timely manner.

Inputs to the current matchup data base are given in Section 3.1.1 in Tables 2, 3, and 4. Contents of the matchup database will be expanded to include data from Brown's *in situ* observation program and the optical moorings deployed by Clark. The external data sets specified by the MOT PI's will be included to correspond with the matchup observations.

3.2.7 Output Product

Output matchup database entries for SST and ocean color for this product are vectors composed of the retrieved SST value, input calibrated radiances and derived brightness temperatures for each channel, ancillary data, flags which quantify the cloud screening results, latitude, longitude and time. A description of the vector components and data types is given in Tables 2 and 3 in Section 3.1.1. These entries will be updated to reflect *in situ* and ancillary data either collected or required by the MOT PI's. This update will begin with receipt of the ATBDs from the MOT PI's.

4.0 Constraints, Limitations, Assumptions

The matchup database will contain data for both "good" and "bad" retrieval times, locations. No a priori assumptions will be made as to use of quality flags for retention of samples within the database. This approach is taken since the data quality criteria will likely change as the satellite retrieval algorithms evolve. Algorithm constraints will be determined by the individual MOT PI's.

5.0 REFERENCES

- André, J.-M. and A. Morel, Atmospheric corrections and interpretation of marine radiances in CZCS imagery, revisited, *Oceanologica Acta*, **14**, 3-22, 1991.
- Austin, R. W. and T.J. Petzold, Remote sensing of the diffuse attenuation coefficient of sea water using the Coastal Zone Color Scanner, in *Oceanography from space*, edited by J. R. F. Gower, pp. 239-256, Plenum Press, New York, 1981.
- Ball Aerospace Division, Boulder CO, Development of the Coastal Zone Color Scanner for Nimbus-7, Volume 2 -- Test and performance data, Final Report F78-11, Rev. A NASA Contract NAS5-20900, May 1979.
- Bowman, K. P. and A. J. Krueger, A global climatology of total ozone from the Nimbus 7 total ozone mapping spectrometer, *J. Geophys. Res.*, **90**(D5), 7967-7976, 1985.
- Bricaud, A. and A. Morel, Atmospheric corrections and interpretation of marine radiances in CZCS imagery: Use of a reflectance model, *Oceanologica Acta*, **7**, 33-50, 1987.
- Clark, D. K., Phytoplankton algorithms for the Nimbus-7 CZCS, in *Oceanography from Space*, edited by J. R. F. Gower, pp. 227-238, Plenum Press, New York, NY, 1981.
- Clark L.A. and D.P Pregibon. 1991. Tree-Based Models. [In] Statistical Models in S, Wadsworth & Brook/Cole, p.377.
- Cox, C. and W. Munk, Measurements of the roughness of the sea surface from photographs of the sun's glitter, *J. Opt. Soc.Am.*, **44**, 838-850, 1954.
- Evans, R. H. and H. R. Gordon, CZCS "System Calibration:" A retrospective examination, Submitted to the *Journal of Geophysical Research --- Oceans*
- Feldman, G. C., N. Kuring, C. Ng, W. Esaias, C. R. McClain, J. Elrod, N. Maynard, D. Endres, R. Evans, J. Brown, S. Walsh, M. Carle and G. Podestá. Ocean Color: Availability of the Global Data Set, *EOS Trans. Amer. Geophys. Union*, **70**, 634-641, 1989.
- Fraser, R. S. and Y. J. Kaufman, Calibration of satellite sensors after launch, *Appl. Optics*, **25**, 1177-1185, 1986.
- Gordon, H. R., Removal of atmospheric effects from satellite imagery of the oceans, *Appl. Optics*, **17**, 1631-1636, 1978.

- Gordon, H. R., A preliminary assessment of the Nimbus-7 CZCS atmospheric correction algorithm in a horizontally inhomogeneous atmosphere, in *Oceanography from Space*, edited by J. R. F. Gower, pp. 257-266, Plenum Press, New York, NY, 1981a.
- Gordon, H. R., Reduction of error introduced in the processing of coastal zone color scanner-type imagery resulting from sensor calibration and solar irradiance uncertainty, *Appl. Optics*, **20**, 207-210, 1981b.
- Gordon H. R., Calibration requirements and methodology for remote sensors viewing the oceans in the visible, *Remote Sens. Environ*, **22**, 103-126, 1987.
- Gordon, H. R., Radiometric considerations for ocean color remote sensors, *Appl. Optics*, **29**, 3228-3236, 1990.
- Gordon, H. R., Some studies of atmospheric optical variability in relation to CZCS atmospheric correction, NOAA National Environmental Satellite and Data Information Service, Final Report Contract No. NA-79-SAC-00714, February 1984.
- Gordon, H. R., J. W. Brown, O. B. Brown, R. H. Evans and D. K. Clark, Nimbus-7 Coastal Zone Color Scanner: reduction of its radiometric sensitivity with time, *Appl. Optics*, **22**, 3929-3931, 1983.
- Gordon, H. R., J. W. Brown and R. H. Evans, Exact Rayleigh scattering calculations for use with the Nimbus-7 Coastal Zone Color Scanner, *Appl. Optics*, **27**, 862-871, 1988.
- Gordon, H. R., O. B. Brown, R. H. Evans, J. W. Brown, R. C. Smith, K. S. Baker and D. K. Clark, A semi-analytic radiance model of ocean color, *J. Geophys. Res.*, **93D**, 10909-10924, 1988.
- Gordon, H. R. and D. J. Castaño, The Coastal Zone Color Scanner atmospheric correction algorithm: Multiple scattering effects, *Appl. Optics*, **26**, 2111-2122, 1987.
- Gordon, H. R. and D. K. Clark, Atmospheric effects in the remote sensing of phytoplankton pigments, *Boundary-Layer Meteorology*, **18**, 299-313, 1980.
- Gordon, H. R. and D. K. Clark, Clear water radiances for atmospheric correction of Coastal Zone Color Scanner imagery, *Appl. Optics*, **20**, 4175-4180, 1981.
- Gordon, H. R., D. K. Clark, J. W. Brown, O. B. Brown, R. H. Evans and W. W. Broenkow, Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison between *in situ* determinations and Coastal Zone Color Scanner estimates, *Appl. Optics*, **22**, 20-36, 1983.

- Gordon, H. R., D. K. Clark, J. L. Mueller and W. A. Hovis, Phytoplankton pigments derived from the Nimbus-7 CZCS: initial Comparisons with surface measurements, *Science*, **210**, 63-66, 1980.
- Gordon, H. R. and A. Y. Morel, *Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review*, Springer-Verlag, New York, 1983, 114 pp.
- Gordon, H. R. and M. Wang, Surface roughness considerations for atmospheric correction of ocean color sensors. 1: The Rayleigh scattering component, *Appl. Optics*, **31**, 4247-4260, 1992a.
- Gordon, H. R. and M. Wang, Surface roughness considerations for atmospheric correction of ocean color sensors. 2: Error in the retrieved water-leaving radiance, *Appl. Optics*, **31**, 4261-4267, 1992b.
- Hansen, J. E. and L. D. Travis, Light scattering in planetary atmospheres, *Space Science Reviews*, **16**, 527-610, 1974
- Hofmann, D. J., Aerosols from past and present volcanic emissions, in *Aerosols and Climate*, edited by P. V. Hobbs and M. P. McCormick, pp. 195-214, A. Deepak Publishing, Hampton, VA, 1988.
- Hooker, S. B., W. E. Esaias, G. C. Feldman, W. W. Gregg and C. R. McClain, *SeaWiFS Technical Report Series: Volume 1, An overview of Sea WiFS and Ocean Color*, NASA Technical Memorandum 104566, July 1992.
- Hovis, W. A., D. K. Clark, F. Anderson, R. W. Austin, W. H. Wilson, E. T. Baker, D. Ball, H. R. Gordon, J. L. Mueller, S. Y. E. Sayed, B. Strum, R. C. Wrigley and C. S. Yentsch, Nimbus 7 coastal zone color scanner: System description and initial imagery. *Science*, **210**, 60-63, 1980.
- Hovis, W. A., J. S. Knoll and G. R. Smith, Aircraft measurements for calibration of an orbiting spacecraft sensor, *Appl. Optics*, **24**, 407-410, 1985.
- Kenizys, F. X., E. P. Shettle, W. O. Gallery, J. H. Chetwynd, L. W. Abreu, J. E. A. Selby, S. A. Clough and R. W. Fenn, Atmospheric Transmittance/Radiance: The LOWTRAN 6 Model, Air Force Geophysics Laboratory, Hanscomb AFB, MA 01731, AFGL-TR-83-0187, 1983, NTIS AD A137796.
- Koepke, P., Vicarious satellite calibration in the solar spectral range by means of calculated radiances and its application to Meteosat, *Appl. Optics*, **21**, 2845-2854, 1982.

- Mie, G., Beiträge zur Optik trüber Medien, speziell kollidalen Metall-lösungen, *Ann. Phys.*, **25**, 377-445, 1908.
- Minnett, P.J. 1991. Consequences of seasurface temperature variability on the validation and applications of satellite measurements. *Journal of Geophysical Research* **96**:18475-18489
- Mueller, J. L., Nimbus-7 CZCS: Confirmation of its radiometric sensitivity decay rate through 1982, *Appl. Optics*, **24**, 1043-1047, 1985.
- Mueller, J. L., Nimbus-7 CZCS: Electronic overshoot due to cloud reflectance, *Appl. Optics*, **27**, 438-440, 1988
- Neckel, H. and D. Labs, The solar radiation between 3300 and 12500 Å, *Solar Phys.*, **90**, 205-258, 1984.
- Salomonson, V. V., W. L. Barnes, P. W. Maymon, H. E. Montgomery and H. Ostrow, MODIS: Advanced facility instrument for studies of the earth as a system, *IEEE Geosci. Rem. Sens.*, **27**, 145-152, 1989.
- Shettle, E. P. and R. W. Fenn, Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties, Air Force Geophysics Laboratory, Hanscomb AFB, MA 01731, AFGL-TR-79-0214, 1979.
- Slater, P. N., S. F. Biggar, R. G. Holm, R. D. Jackson, Y. Mao, M. S. Moran, J. M. Palmer and B. Yuan, Reflectance- and radiance-based methods for the in-flight absolute calibration of multispectral sensors, *Remote Sens. Environ.*, **22**, 11-37, 1987.
- Smith, R. C. and W. H. Wilson, Ship and satellite bio-optical research in the California Bight, in *Oceanography from Space*, edited by J. F. R. Gower, pp. 281-294, Plenum, New York, NY, 1981.
- Vigroux, E., Contribution à l'étude expérimentale de l'absorption de l'ozone, *Ann. Phys. Paris*, **8**, 709-762, 1953.
- Viollier, M., Radiance calibration of the Coastal Zone Color Scanner: A proposed adjustment, *Appl. Optics*, **21**, 1142-1145, 1982.
- Viollier, M. and B. Sturm, CZCS data analysis in turbid coastal water, *J. Geophys. Res.*, **89**(D4), 4977-4985, 1984.
- Wang, M., Atmospheric Correction of the Second Generation Ocean Color Sensors, 1991, Ph.D. Dissertation, University of Miami, Coral Gables, FL, 135pp.

Appendix 1 Equal-area gridding scheme

Equal-area gridding scheme implemented for MODIS

A1.1 Introduction

This document describes the equal-area gridding scheme proposed by the RSMAS Remote Sensing Group for the binned ocean fields produced for MODIS. The same approach has been adopted for AVHRR Ocean Pathfinder SST and SeaWiFS binned ocean color products. The gridding scheme is based on that adopted by the International Satellite Cloud Climatology Project (ISSCP).

This document does not motivate the need for an equal area grid for SeaWiFS or other oceanographic products. Such motivation can be found in a paper by W. Rossow and L. Gardner (Selection of a map grid for data analysis and archival, *Journal of Climate and Applied Meteorology*, 1984, 23:1253-1257). Furthermore, this document describes only the design of the proposed equal-area grid, and does not discuss other related topics such as rules for spatially or temporally combining observations into the equal-area bins. These considerations are presented in Appendix 3.

A1.2 Overview

The gridding scheme proposed consists of rectangular bins or tiles, arranged in zonal rows. A compromise between data processing and storage capabilities, on one side, and the potential geophysical applications of satellite data, on the other side, suggest that a suitable minimum bin size would be approximately 8-10 km on a side.

In the scheme proposed here, the tiles are approximately 9.28 km on a side and can accommodate bin sizes of 1 and 4 km as well. This size (9.28 km) was chosen because (a) it has approximately the desired minimum resolution, and (b) it results in 2160 zonal rows of tiles from pole to pole (i.e., 1080 in each hemisphere). This particular number of rows (2160) has some advantages which will be discussed in more detail below. Because the total number of rows is even, the bins will never straddle the Equator (i.e., there will

be an equal number of rows above and below the Equator). This avoids possible situations where the Coriolis factor is zero, a characteristic that numerical modellers expect from any gridding scheme adopted.

The total number of approximately 9-km bins is 5,940,422. The bins or tiles are arranged in a series of zonal rows; the number of tiles per row varies. The rows immediately above and below the Equator have 4320 tiles. This number is derived by dividing the perimeter of the Earth at the Equator by the standard tile size (i.e., $2 R_e/9.28$), where R_e is the equatorial radius of the Earth ($R_e = 6378.145$ km). The number of tiles per row decreases approximately as a cosine function as the rows get closer to each pole (rigorously, there should be an adjustment for ellipticity of the Earth, as the equatorial radius decreases progressively to the smaller polar radius; this adjustment is not applied in the current implementation). At the poles, the number of tiles is always three. This special situation will be discussed in detail below. The number of tiles per row as a function of latitude is shown on Figure A1-1.

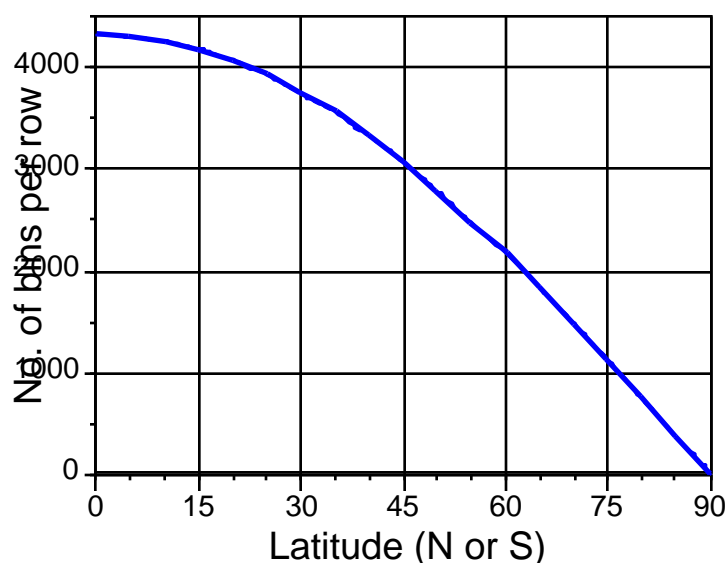


Figure A1-1. Number of 9.28 km tiles per zonal row as a function of latitude (North or South). The number of tiles is 4320 at the Equator and decreases to 3 at the poles.

The number of bins in each zonal row is always an integer. To ensure an integer number of bins, the width of each bin (the size of a bin along a parallel, or x-length) must vary slightly from row to row. The bins, however, are always 9.28 km long along the meridians. That is, only one of the bin dimensions changes. The size of the bins at

each zonal row is established in the following manner. First, a preliminary value for the number of tiles (N_p) at a given latitude (L) is computed as

$$N_p = 2 \ r / X,$$

where X is the x-size of a bin at the Equator (9.28 km) and r is the radius of the circle produced by slicing the Earth with a plane parallel to the Equator at latitude L . The radius r can be calculated as

$$r = R_e \cos(L),$$

where R_e is the equatorial radius of the Earth. If the fractional part of N_p is greater or equal than 0.5, then N_p is rounded up to the nearest integer (i.e., the final number of tiles will be the integer portion of N_p plus one), otherwise N_p is rounded down (the final number of tiles is the integer portion of N_p). Once the final integer number of tiles along a row is calculated, the X -size of the tiles must be adjusted. This is done by dividing the perimeter of the row ($2 \ r$) by the integer number of tiles. The result is the x-length of a tile (width) for a given row.

Because the x-length of the tiles is adjusted to ensure an integer number at each row, the “equal area” characteristics of this binning scheme are not rigorously preserved. However, variations in tile size are negligible throughout most of the globe, and only become relevant at very high latitudes, where there are fewer tiles per row and, thus, any adjustments are more noticeable. As soon as the number of tiles increases with distance from the poles, the difference between tile sizes rapidly becomes practically unnoticeable. To provide an idea of the magnitude of the fluctuations in tile size, the worst possible case occurs when half a tile remains “uncovered” after filling a zonal row with an integer number of tiles. Once a row has 100 bins (approximately 16 rows, or 148 km from the poles), the worst possible difference between the actual tile x-length and the standard x-length is of the order of 0.5% (i.e., half a tile's length redistributed among about 100 tiles). For a tile of about 9 km a side, this represents a difference in the x-length of about 45 m. Through a similar calculation, a row with 50 bins (about 80 km away from the poles) has a 1% variation with respect to the standard bin size.

The gridding scheme described here has an extremely useful feature: the number of 9.28 km tiles in each hemisphere (1080) is divisible by many numbers (e.g., 2,3,4,5,6) and therefore it is extremely easy to generate an integer number of rows at many useful

spatial resolutions. For instance, 12 rows of 9.28 km tiles can be combined to generate zonal bands of approximately one degree (one degree of latitude is equal to 111.12 km; 12 bins would form a band 111.20 km wide). Another example is the use of 30 rows of to generate zonal bands of approximately 2.5° (a typical output resolution of atmospheric circulation models).

A1.3 The poles

Both the North and South poles are special cases in the gridding scheme presented here. The pole areas are always covered by three tiles, shaped like pie sectors. While the meridional size of the polar bins (the y-length) will be the usual 9.28 km, the length of the bins along the arc of the sectors will be slightly larger. Neglecting sphericity, the area encompassed by the last row of tiles is X^2 , where $X = 9.28$ km. If we express the area of the circle as a rectangle of height X , the remaining dimension is X . If we divide the perimeter by three (to yield three tiles), each tile will have dimensions X by $X/3$ (approximately 1.05 X). That is, the bases of the triangular polar tiles are about 5% larger than the x-length of the equatorial tiles.

A1.4 Binning software

Several routines have been developed to perform the principal transformations required for binning and mapping data, such as converting latitudes and longitudes into bin numbers. Other routines perform the inverse transformation, that is, given a bin number they return a latitude and longitude corresponding to the centroid of that bin. These routines use a common initialization routine that must be executed prior to calling the conversion routines.

Appendix 2 Data Day Definition

A 2.1 Introduction

The basic products generated by both the AVHRR Pathfinder and the SeaWiFS projects are global daily fields of geophysical quantities such as sea surface temperature and chlorophyll concentration. This definition is proposed for use with global fields generated from MODIS products. The daily fields will be the basis of subsequent temporal compositing into weekly and monthly products. One basic question, however, is: what constitutes “a day’s worth” of data? This is the question we address in this document.

The need for a consistent definition of a data-day is only really relevant for the production or analysis of *global* data fields. If one is dealing with a limited area (although, in this case, “limited” means anything less than global, and can encompass entire ocean basins), one takes advantage of the fact that the satellite sensors usually sample a region at *approximately* the same time(s) every day. In this way, data separated by approximately 24-hour periods can be assigned to different data-days (a further separation into daytime and nighttime fields can be made with the AVHRR). Analyses of the resulting daily data fields will introduce a minimal amount of temporal aliasing, as the difference in sampling times is of the order of a couple of hours over an approximate repeat cycle of a few days.

In contrast, when daily global satellite data fields are to be constructed, a consistent definition of a data-day needs to be adopted. This definition should be easy to implement in practice and should minimize temporal aliasing and discontinuities in the resulting products. In the following paragraphs we explore some of the alternatives.

A2.2 A 24-hour data day

The most obvious definition of a data-day is a 24-hour period. For instance, a daily field would encompass all the data collected between 00:00:00 UTC (or any other arbitrary start of the day) and 23:59:59 UTC. This definition is simple, intuitive, and extremely easy to implement. Its negative aspects, however, become apparent when one considers the orbital characteristics of the spacecraft on which the sensors of interest are, or will be, flown.

To illustrate the problem, we present a plot of nadir tracks for the NOAA-11 spacecraft (Figure A2-1). To simplify the visualization we only display descending tracks (i.e., the spacecraft is flying from north to south). The NOAA descending tracks correspond to nighttime data, although, in the case of SeaWiFS, the descending tracks will correspond to the daytime data (the only data archived for this sensor, other than special calibration measurements).

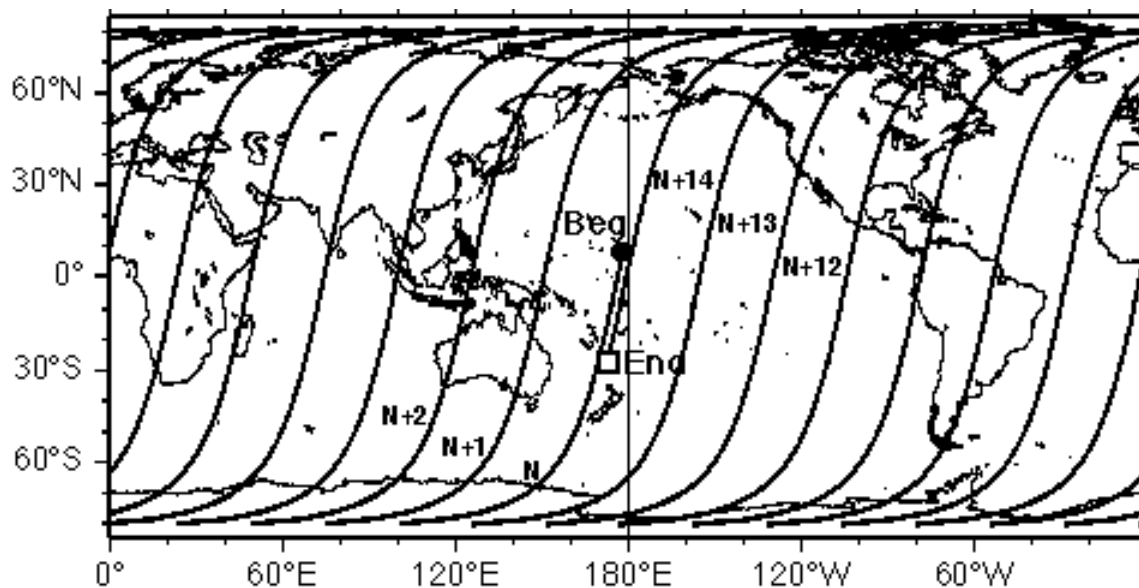


Figure A2-1. Descending NOAA-11 tracks for a 24-hour data-day beginning on July 26 1992 15:22:00 UTC. The data-day begins at the point labeled “Beg” and ends at the square labeled “End”. The first orbit after the beginning of the data-day is labeled “N”, and subsequent orbits are “N+1”...“N+4”.

For comparison with subsequent cases, we choose to begin the 24-hour data-day on July 26, 1992 at 15:22:00 UTC, when the nadir track intersects the 180° meridian (Marked

“Beg” on Figure A2-1). The descending orbit immediately after the beginning of the data-day is labeled N. Subsequent descending tracks pass to the west, and are offset by a distance of about 25 degrees of longitude at the Equator. The swaths viewed by the AVHRR in consecutive orbits have an increasingly larger overlap with latitude. This means that areas at intermediate and high latitudes may be sampled twice or more during a data-day (ignoring, for the time being, the ascending orbits). When an area is sampled in two consecutive descending orbits, measurements will be separated by about an hour and a half. Unless one is concerned with features with very small scales, it is probably safe to assume that the ocean fields will not change significantly between consecutive passes, thus temporal aliasing should be negligible.

The NOAA polar platforms that carry the AVHRR have an orbital period of approximately 102 minutes. The actual period depends on the spacecraft altitude, and, therefore, will be slightly different for each NOAA spacecraft. The orbital period may also vary with time, as the altitude of a satellite changes. Given an orbital period of about 102 minutes, the number of revolutions that the NOAA spacecraft will complete in a 24-hour period is approximately 14.12. The last descending orbit of the 24-hour data day is labeled N+14. It is apparent from Figure A3-1 that there is along-track overlap between descending tracks N and N+14. The areas in which there is overlap will have been sampled twice (ignoring the smaller overlap between consecutive orbits) in a data-day, and the output will be the average of measurements taken almost 24 hours apart. The 24-hour definition of a data-day, therefore, may result in temporal aliasing in areas near the beginning and end of the 24-hour period, due to the inclusion in a given day of overlapping orbit tracks sampled almost 24 hours apart.

A second problem of the temporal definition of a data-day is the existence of areas on the global fields with large temporal discontinuities in sampling times, even though they may be spatially contiguous. For instance, in Figure A2-1, one can see descending track N+14, the last track of the data-day. To the north of that track (i.e., over the Arctic Ocean north of Alaska), data are contributed by track N+1 and, possibly, N+2. These two tracks, however, were sampled near the beginning of the data-day, more than 20 hours before track N+14. The daily fields, then, will contain large temporal discontinuities along the boundaries between data swaths from tracks N+14 and N+1. If there is overlap between the two swaths, data collected far apart in time may be averaged, once again introducing potential aliasing. Similar problems occur in the area

south of track N (south of New Zealand), which is sampled by tracks N+13 and N+12 much later in the day.

The aliasing and temporal discontinuity effects are further complicated by the fact that the locations where they occur change in time. Figure A2-2 shows the locations along the nadir tracks of the boundaries between 24-hour data-days for a 10-day period beginning on July 26, 1992 (for the NOAA-11 spacecraft). The dot labeled “1” corresponds to the beginning of the cycle on July 26, 1992 at 15:22:00 UTC. The dot labeled “2” indicates the beginning of the second 24-hour data-day, and so forth. The shift in the location of the daily boundaries is a direct result of the difference between the 24-hour data-day and the shorter time it takes the spacecraft to complete a number of revolutions that would ensure global coverage.

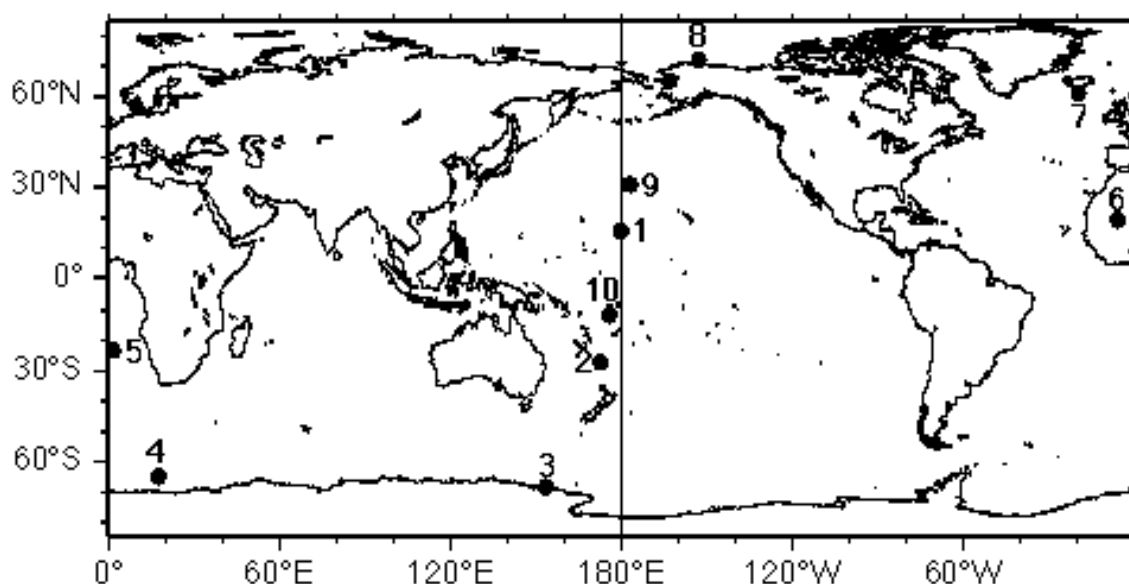


Figure A2-2. Locations of the boundaries of 24-hour data days for a 10-day period beginning on July 26, 1992 15:22:00 UTC (dot labeled “1”).

A2.3 A spatial data-day definition

Because of the problems associated with a temporal definition of a data-day, we explored the implications of adopting a spatial definition. In this case, the boundary between data-days is not defined by time but, instead, by a fixed geographic reference. A similar criterion is commonly used for designating orbit numbers in several

spacecraft: the orbit number usually is incremented upon crossing the Equator. For the initial investigations, we selected the 180° meridian as the boundary between data-days.

Figure A2-3 shows NOAA-11 nadir tracks for a spatially-defined data day. Because the nadir tracks crosses the reference line several times during a day, one of the crossings must be selected as the beginning of a data-day. An operational definition of this is presented below. For this discussion, we define the day to begin on July 26, 1992 at 15:22:00 UTC, when the spacecraft crosses the 180° meridian flying from north to south (i.e., at the same time at which the 24-hour data-day shown on Figure A2-1 started). The first descending track of the day is labeled N.

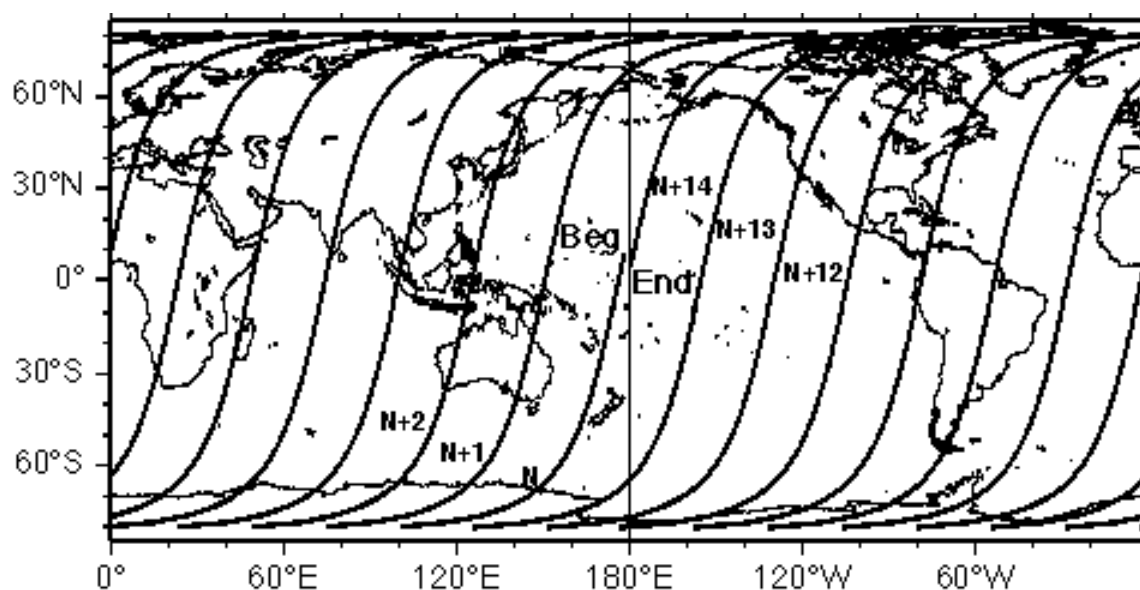


Figure A2-3. NOAA-11 descending orbits for a spatially-defined data-day beginning on July 26 1992, 15:22:00 UTC. At this time, the nadir track crosses the 180° meridian.

As almost 24 hours worth of data are required to ensure global coverage, we define the data-day as ending when the nadir track crosses the 180° meridian during revolution N+14. This happens approximately on June 27, 1992 at 15:14:00 UTC. The most immediate observation, then, is that a spatial definition results in a data-day that does not necessarily correspond to a 24-hour day: in this case the data-day is approximately 23 hours and 52 minutes long. This figure is only approximate for two reasons. In the first place, it is sometimes necessary to include an additional revolution in order to ensure global coverage (that is, the last orbit of the day would be N+15). Secondly, the

spatial definition is applied on a pixel-by-pixel basis. That is, pixels along the same scan line on a given orbit can be assigned to different days depending on whether they are on one side or the other of the 180° meridian.

Figure A2-4 illustrates the pixel-by-pixel assignment of data to a given day. The figure shows the sampling pattern of the AVHRR onboard NOAA-11 between 15:12:00 and 15:32:00 UTC on July 26, 1992 (i.e., ± 10 minutes from the start of the data-day at 15:22:00 UTC). The scan lines shown on Figure A2-4 are separated by one minute (in a one-minute interval there are 360 LAC scans or 120 GAC scans). Pixels along a given scan line that are located east of 180° are assigned to day N. If pixels along the same scan line are west of 180°, those pixels are assigned to the following day (N+1). It is apparent from Figure A2-4 that, even before the nadir track crosses the 180° meridian, pixels are already being assigned to day N+1. Conversely, after the nadir track has crossed the reference meridian (at 15:22:00 UTC), pixels east of the meridian are still being allocated to day N. It is this allocation mechanism that makes it difficult to define precisely the duration of a data-day.

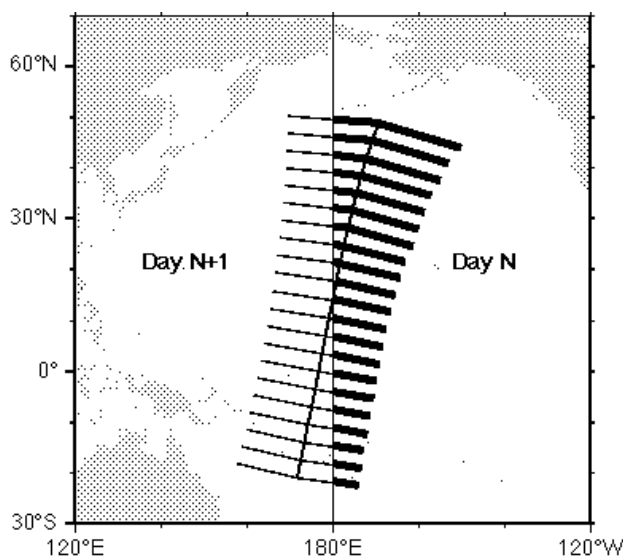


Figure A2-4. AVHRR nadir track and scan lines for a 20-minute period between 15:12:00 and 15:32:00 UTC on July 26, 1992. Pixels to the east of the 180° meridian (marked in a thicker line) get assigned to data-day N, whereas the pixels to the west of the meridian correspond to data-day N+1.

A2.4 How is the beginning of a data-day defined?

How is the spatial definition of a data-day implemented in routine processing of global satellite data fields? The first step is to define a meridian which will serve as the reference for the data-day definition. The 180° meridian used in the previous examples is a good alternative, as this choice minimizes differences between actual dates and the dates assigned to the data-days. As the spatial data-days are not 24-hours long, a suitable naming convention will have to be established.

A second step in defining a data-day is to decide which of the descending (or ascending) crossings of the reference meridian will mark the beginning of the descending (or ascending) data-days. As mentioned above, there are several (usually seven to nine) descending crossings of the reference meridian in a day; the same is true for ascending orbits. This is illustrated in Figure A2-5, which shows the latitude of descending crossings of the 180° meridian as a function of time for the NOAA-11 spacecraft, beginning on July 26, 1992; a 10-day span is shown. Most of the crossings (shown as dots) take place at high latitudes, and one or two crossings per day occur at tropical to intermediate latitudes.

Any of the crossings of the 180° meridian shown on Figure A2-5 can be potentially selected as the one marking the beginning of a data-day for descending and ascending orbits. For operational purposes, we propose the following definition: *a data-day for descending orbits is defined to begin at the descending crossing of the 180° meridian closest to the Equator.* A similar definition can be applied to ascending crossings, yielding data-days for ascending orbits. Such definition is the easiest to implement because there is always only one crossing in a day that fulfills the condition (although consecutive crossings may sometimes have very similar absolute latitudes of intersection, one on the Southern hemisphere, and the other on the Northern hemisphere).

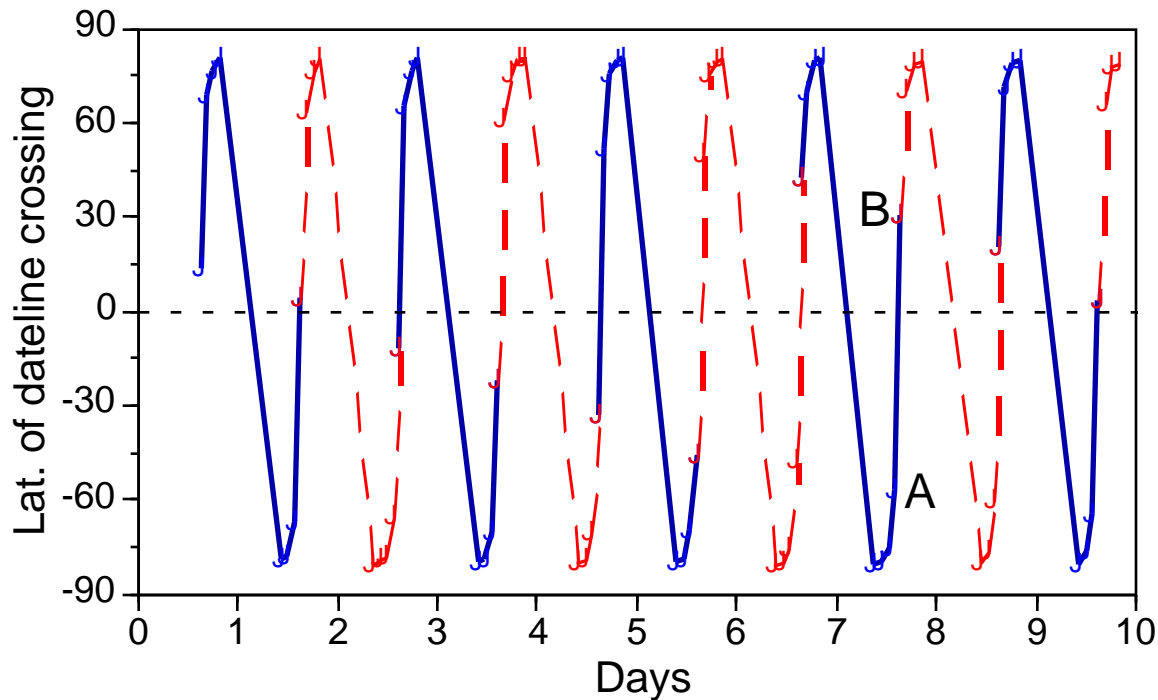


Figure A2-5. Latitude of crossing of the 180° meridian for NOAA-11 descending orbits. Data shown for a 10-day period beginning on 26 July 1992, 15:22:00 UTC. The alternating solid and dashed lines indicate consecutive data-days.

The alternating solid and dashed lines in Figure A2-5 indicate consecutive data-days. Initially, the latitude of the data-day beginning seems to follow a regular progression to the south. For instance, the first two data-day boundaries in Figure A2-5 are on the Northern hemisphere, and the next four are progressively further south on the Southern hemisphere. Note, however, that the southward progression is interrupted near the end of data-day 6 (the point labeled A). In this case, the next descending crossing (point labeled B) is actually closer to the Equator, so the data-day is extended until this next crossing (located in the Northern hemisphere). That is, the data-day is slightly longer (one more revolution) in this case. The southward progression of the crossings subsequently resumes. For ascending orbits, the progression of the latitude of the crossings is reversed, that is, it occurs from south to north.

Table A2-1 contains a list of start times of descending data-days for a 15-day period beginning on July 26, 1992, as well as the latitude at which the crossing of the 180° meridian occurs.

Table A2-1. Beginning times of fifteen data-days for descending orbits, NOAA-11 spacecraft. The latitude of the 180° meridian crossing is also shown.

Date	Beginning time	Latitude of 180° crossing
07/26/92	15:22:04	13.85
07/27/92	15:13:39	0.67
07/28/92	15:05:14	-12.52
07/29/92	14:56:34	-24.81
07/30/92	14:47:28	-35.52
07/31/92	14:37:51	-44.37
08/01/92	15:43:42	39.33
08/02/92	15:34:24	29.34
08/03/92	15:25:35	17.60
08/04/92	15:17:06	4.63
08/05/92	15:08:42	-8.64
08/06/92	15:00:08	-21.29
08/07/92	14:51:11	-32.58
08/08/92	14:41:43	-41.99
08/09/92	15:47:31	41.94

We must stress that, because of the pixel-by-pixel allocation described above, parts of the field will include data collected both before and after the times listed in Table A2-1. Notice the jump in the southward progression of crossing latitudes (e.g., from July 31 to August 1), which is associated with a slightly longer data-day.

A2.5 Advantages of the Spatial Definition of a Data-Day

In previous sections we proposed a spatial definition for a data-day, together with an objective definition for the temporal “beginning” and “end” of such a data-day. So far, however, we have not discussed the advantages or disadvantages of the proposed definitions.

Some problems associated with a temporal definition of the data-day were the potential presence of aliasing and large temporal discontinuities, and the fact that the day boundaries changed with time. The spatial definition avoids temporal changes in the location of boundaries, as the boundary is fixed (e.g., the 180° meridian). Furthermore, because there is no overlap of swaths at the beginning and end of a data-day, the spatial

definition reduces the aliasing resulting from averaging data sampled almost 24 hours apart. The presence of large temporal discontinuities among adjacent areas is still present, however.

The large temporal discontinuities identified on Figure A2-1 north of Alaska and south of New Zealand are still present in Figure A2-3. It is clear that the large temporal discontinuities occur in two places near the meridian that defines the separation between data-days. The first place is the area south of the first track of the data-day and west of the reference line. The second area with discontinuities occurs north of the last track of the data-day, east of the reference line. In addition to the large temporal discontinuities between adjacent swaths, when the swaths overlap at higher latitudes once again data will be averaged that were sampled far apart in time. Elsewhere on the global fields, any given track is surrounded by tracks sampled one orbital period (about 100 minutes) earlier or later.

The presence of temporal discontinuities or the averaging of data collected at very different times may not be too important for many applications, although users should certainly be made aware of the occurrence of these events. In other situations, however, the temporal discontinuities may cause significant problems. Examples of such applications may be the estimation of the translation speed of certain features, or the computation of fluxes.

In order to limit the large meridional temporal discontinuities near the data-day boundary, the short track segments north and south of the first and last tracks of the data-day could simply be eliminated (e.g., parts of N+1, N+2, N+12 and N+13). This approach is illustrated in Figure A2-6, which shows descending tracks between July 27 1992, 15:14:00 UTC and July 28 1992, 15:05:00 UTC (the data-day following the one shown on Figure A3-3). The map is now centered at 0°, rather than at 180° as in Figure A2-3. Note that the nadir tracks for which segments were eliminated seem to end a bit before or after the 180° line. This is because positions were predicted at one-minute increments.

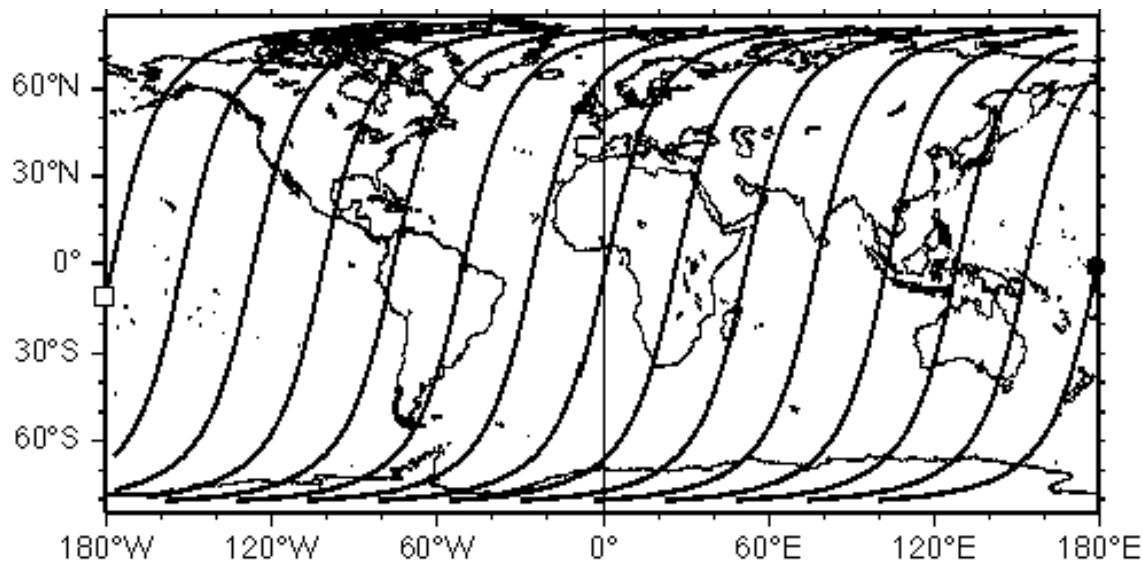


Figure A2-6. NOAA-11 descending orbits for spatially-defined data-day beginning on July 27, 1992 15:14:00 UTC. Segments that introduce large north-south temporal discontinuities (see text) are excluded.

The elimination of segments may result in areas not being sampled (e.g., upper left and lower right corners of the map). These gaps may possibly be filled by the wide swath of the first and last tracks of the data-day (tracks N+14 and N in the north and south, respectively). However, the size of the gaps is a function of the latitude of the reference line crossing which defines the beginning of the data-day. As shown above (Figure A2-5), this latitude changes with time, moving north and south approximately between 60°N and 60°S. When the crossing is further north, the gap to the south of the first track will be larger. Conversely, when the crossing is further south, the gap north of the last track will get larger. We propose that two additional swaths be added at each end of the data-day, in order to replace the eliminated segments. Experience has shown that two additional swaths are enough to fill each of the gaps and ensure complete coverage. The added swaths would be temporally continuous with the first and last tracks of each data-day, thus eliminating the problems of temporal discontinuities. An operational scheme would involve the following steps:

1. The times of the beginning and end of a spatially-defined data-day are found following the definition suggested above. These times will be referred to as the “beginning” and “end” of the data-day.
2. Data east of the 180° meridian and collected up to 216 minutes (about two orbits) after the beginning of the data day will be excluded. Data west of the 180° meridian

sampled up to 216 minutes before the end of the data-day will be similarly excluded. The net result of these actions is similar to the elimination of segments shown on Figure A2-6.

3. To ensure full coverage, data collected up to 216 minutes before the beginning of the data-day and west of the 180° meridian are added to the data-day. This fills the gap to the south of the first track of the day. Data collected up to 216 minutes after the end of the data-day and east of 180° are also added. These data fill the gap north of the last track of the data-day. The end result is illustrated on Figure A2-7. Note that only the descending (AVHRR night or ascending day) portions of the extra orbits are included in the fields.

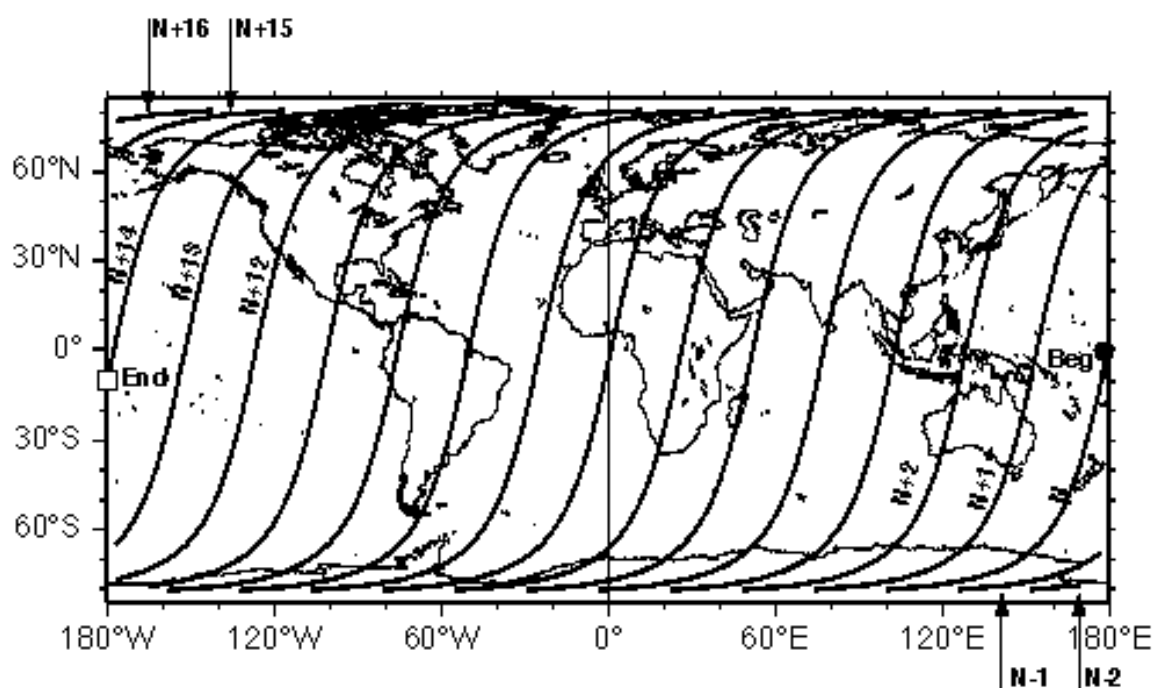


Figure A2-7. Data-day beginning on July 27 1992 15:14:00 UTC, showing the addition of four segments (indicated by arrows) in order to minimize temporal discontinuities. The first track sampled after the estimated beginning time of the day ("Beg") is track N. The two segments to the south correspond to the two previous orbits (N-1, N-2). The last track before the estimated end time of the data-day ("End") is track N+14. The two segments to the north correspond to the next two orbits (N+15, N+16).

Figure A2-7 shows the descending orbits for the data-day beginning approximately on July 27 1992, 15:14:00 UTC. The gaps shown on Figure A2-6 have been filled by the addition of four short segments, indicated by arrows on Figure A2-7. Note that these

segments have been sampled before (N-1, N-2) and after (N+15, N+16) the times estimated for the beginning and end of this data-day (see Table A2-1). However, because the added segments are close in time to orbits N and N+14, the large temporal discontinuities have been eliminated. The segments excluded from this data-day are the first portion of tracks N+1 and N+2, east of 180°, and the last portion of tracks N+12 and N+13, west of 180°.

Admittedly, it is somewhat difficult to grasp the methodology proposed. To facilitate comprehension, we may present a simple analogy. Envision a continuous strip chart on which the continents are drawn. Above the chart recorder there is a clock showing UTC time and date. As the chart moves from left to right, a pen draws descending tracks, one at a time. The speed of the chart movement is appropriate to ensure that the nadir track's latitude and longitude corresponding to any given UTC time are correct. That is, the nadir tracks should look similar to those on Figures A2-6 and A2-7.

Suppose we position the chart so that the pen is just crossing the 180° meridian near the Equator on July 27 1992. The time shown by the clock should be about 15:14:00 UTC. We then allow the chart recorder to run for almost 24 hours, until a track crosses the 180° meridian again very close to the Equator. The time should be about 15:04:00 UTC on July 28, 1992. If we cut the chart along the two 180° meridians drawn (left and right), the tracks on the chart should look exactly like Figure A2-6. As in Figure A2-6, there will be some gaps in the coverage. On the right side of the chart, there is a gap south of the first track (N) of the day. This gap should have been filled by the last portion of tracks N+12 and N+13, which have been drawn to the left of the 180° meridian on the left side of the chart. These lines, however, were eliminated when we cut the chart along the left 180° line. Similarly, the gap north of the last track of the day should have been filled by the initial portions of tracks N+1 and N+2. These segments were drawn east of the 180° meridian on the right side of the plot. However, as we cut along the 180° line on the right, these segments were excluded. It is apparent, then, that the chart recorder analogy reproduces the action of eliminating tracks which cause large temporal discontinuities, as the end result looks exactly like Figure A2-6. Let us see if we can fill the gaps in the global fields using the same chart recorder analogy.

Suppose that we do not start the chart recorder at 15:14:00 UTC on July 27 1992 but, rather, we move the chart backwards and start about 216 minutes earlier. If we start the recorder then, a few additional tracks (e.g., N-1 and N-2) will be drawn before the nadir

track of orbit N crosses the 180° meridian at 15:14:00 UTC (defined as the temporal beginning of the data-day). The southern segments of tracks N-1 and N-2 will fall west of the 180° meridian, filling the gap previously existing in the south. We let the recorder run up to 216 minutes past the time estimated for the end of the day (July 28 1992, 15:04:00 UTC) and, again, a few additional tracks will be drawn. If the last track of the day is N+14, the northern portions of tracks N+15 and N+16 will fill the northern gap. Once we have allowed the recorder to run for about 24 hours plus the additional 216 minutes on either end, we take a pair of scissors and cut the chart along both 180° meridians. That is, we are applying the spatial pixel-by-pixel assignment of data to a given data-day. The end result should look exactly like Figure A2-7. Finally, we could envision running the recorder for long periods and repeatedly cutting the long chart along the 180° meridians. Each of the maps would correspond to one data-day.

When discussing the elimination of segments that caused large temporal discontinuities, we could have given the impression that the data in these segments would be unused, and therefore wasted. However, if one follows the analogy presented above, it is easy to see that the data will not be deleted but, rather, they will be assigned to the previous or the following data-days. For instance, the northern portions of tracks N+1 and N+2 (not labeled) in Figure A2-6 would be plotted to the east of the right 180° meridian on the chart. When we cut the chart, these portions get assigned to the previous data-day, which begins on July 26 1992, 15:22:00 UTC. In the same way, the southernmost portions of tracks N+12 and N+13 are plotted to the west of the left 180° meridian, thus being assigned to the next data-day after the chart is cut along the meridian. The end result of the scheme proposed is a daily global field where all parts of a field are temporally separated from adjacent areas by, at most, one orbital period.

A similar scheme can be implemented for ascending tracks. The definitions of the temporal beginning and end of an ascending data-day were discussed above. The chart recorder analogy can also be similarly formulated for ascending nadir tracks.

A2.6 Other Issues

An aspect that we have not discussed so far is that at both the extreme north and south of the fields, data from several tracks will be averaged for a data-day. At high latitudes,

the spacecraft is flying almost in an east-west direction and, thus, the scan lines have a north-south orientation. For instance, there are seven to nine passes a day at high latitudes (see Figure A2-5). Near the 180° meridian, where some of the passes are excluded at high latitudes, as described above. In other high latitude regions, however, the fields will contain the average of several passes. This should not have too many consequences on ocean products, as the areas affected will be mostly on land in the southern hemisphere and under permanent ice cover in the northern hemisphere.

One final issue that needs to be pointed out is that the spatial scheme proposed above will result in temporal discontinuities in areas that straddle the reference line. Suppose that an oceanographer is studying an area of the North Pacific Ocean encompassed between 150°W and 150°E, straddling the 180° line. If the oceanographer obtains a global field for a given data-day, he/she must realize that the portion of the study area west of 180° has been sampled much earlier than the portion to the east. Again, this may not be relevant for some research, but it could be in some cases. A solution would be to place the reference line elsewhere, for instance along 0°, but there will always be some location where areas on either side of the line will be sampled far apart in time. Alternatively, a user might obtain product fields for two consecutive data-days and paste the appropriate portions. Going back to the Pacific example presented above, the eastern part of the study area would be extracted from data-day X and the western part would be taken from day X+1. This can be accomplished without reprocessing and without introducing any spatial or temporal aliasing.